AD_____

Award Number: DAMD17-97-1-7121

TITLE: Is Mammography Useful in Older Women

PRINCIPAL INVESTIGATOR: Karen Freund, M.D.

CONTRACTING ORGANIZATION: Boston Medical Center
Boston, Massachusetts 02118-2393

REPORT DATE: June 2001

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	June 2001	Final (12 May 97 - 12 May 01)

4. TITLE AND SUBTITLE
Is Mammography Useful in Older Women

5. FUNDING NUMBERS
DAMD17-97-1-7121

6. AUTHOR(S)
Karen Freund, M.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Boston Medical Center
Boston, Massachusetts 02118 -2393
E-MAIL:karen.freund@bmc.org

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT ***(Maximum 200 Words)***

Despite the fact that approximately 43% of all breast cancers occur in women over 70 years of age, no previous randomized controlled trial (RCT) addressed women in this age group. Without clear evidence, the usefulness of the mammograms currently performed on 15-36% of women over age 70 is unknown.

We present the use of an administrative utilization database linked to breast cancer outcomes to examine the usefulness of mammography in women age 70 and older. Using the Linked Medicare-SEER Tumor Registry Database, created by the National Cancer Institute and the Health Care Financing Administration, we have created a database of 9767 women diagnosed with breast cancer from January 1, 1987 to December 31, 1993, and compared mammography users versus nonusers on stage of presentation, all cause survival and breast cancer survival.

Women over 70 years who were nonusers of mammography were diagnosed with breast cancer at Stage II or greater more often than regular users (adjusted odds ratio (OR), 3.12 [95% CI, 2.74-3.58]). Nonusers of mammography were at significantly greater risk of dying from their breast cancer than regular users for all women (adjusted hazard ratio (HR), 3.38 [95% CI, 2.65-4.32]) and for women within each age group. Even assuming a lead-time of 1.25 years, nonusers of mammography continued to be at increased risk of dying from breast cancer.

In order to account for selection bias since we are dealing with observational data, we employed two methods. First, we used propensity score matching to reduce the sample to a simulation of a randomized study. Second, we used instrumental variable analysis to look at unmeasured covariates which are unassociated with region of the country that might affect our outcome. In each case, we showed that there was only a minimal change in the adjusted odds ratio of stage at diagnosis (3.27 and 3.01, respectively). Thus, our results do not appear to be affected by selection bias.

These data support the use of mammography as a screening modality for women 67 years of age and older.

14. SUBJECT TERMS
Breast Cancer, Idea Award

15. NUMBER OF PAGES
58

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

___ Where copyrighted material is quoted, permission has been obtained to use such material.

___ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

___ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

NA In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

X For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

N/A In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

N/A In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

N/A In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Karen Grund 6/4/01

PI - Signature Date

(4) TABLE OF CONTENTS:

(5) INTRODUCTION

Randomized clinical trials

Breast cancer is the second leading cause of cancer-related mortality among women in the United States. Women ages 65 and older bear the greatest burden of disease accounting for more than 43% of newly diagnosed cases of breast cancer (1). Older women are also more commonly diagnosed with advanced stage disease (1-4) and their breast cancer mortality rate is eight times greater than women under age 65 (5). The role of screening mammography in reducing morbidity and mortality from breast cancer in older women is unknown. Randomized controlled trials (RCTs) are inadequate to judge the efficacy of mammography screening, as they did not include sufficient women over age 69 years.

In our last annual report, we presented our methodology to date and data of retrospective cohort study of 9767 women ages 67 and older with breast cancer, diagnosed and staged from 1987 to 1993, in three geographic areas to estimate benefits from prior mammography use for women aged 67-74, 75-84 and 85 and older. Since our last annual report, we have addressed the following technical objectives:

Technical Objective 5:

Task 1

We have developed an innovative methodology, utilizing propensity scores and operational variables to address the issues of the use of observational data to address the benefits of mammography screening in women over age 70.

Task 2

We have published our first manuscript, in the Journal of the American Geriatric Society.

We have submitted our second manuscript describing the statistical methodology to Health Services Outcomes Research for Publication this year. We are presently awaiting their responses and decision on acceptance.

The work on this has won Mr. Posner a student paper award for the Health Policy and Statistics section of the American Statistical Association as well as a first place award at the Boston University Graduate Student Science Day. He will be presenting this work at the Joint Statistical Meetings in Atlanta in August 2001.

(5.1) Background

Mammography Use in Older Women

Early detection with mammography has been consistently shown to decrease breast cancer-related mortality by 30% for women age 50-69 years (6-10). Despite this striking reduction in mortality for women age 50-69, there are currently no data to make a statement about the utility of mammography for women age 70 and older. Although one RCT (The Swedish Two County Trial) included women up to age 74 years, there was inadequate power to detect a difference over age 69 years.

There are reasons to expect that older women would benefit from regular mammography despite the lack of scientific data to establish a benefit. First, mammography is a more specific and sensitive test as women age (11,12). Second, the biology of breast cancer in older women is thought to be similar to women age 50-69 years (13). Third, survival times for older women are sufficiently long to benefit from early detection (14,15). Fourth, the cost effectiveness ratio of breast cancer prevention in the elderly is in a reasonable range (16,17).

Given the lack of scientific data on the usefulness of mammography in women age 70 and older, current practice recommendations vary. Annual mammography is recommended by the American Cancer Society (13) and the American Medical Association Council on Scientific Affairs (18) for women after age 50 with no upper age limit. Annual to biennial mammography is recommended for women age 50-74 by the U.S. Preventive Services Task Force (19). The Task Force does not recommend mammography beyond age 74 (19). Annual mammography is recommended for women age 65-74 by the Forum on Breast Cancer Screening in Older Women. The Forum also suggests that mammography "should be encouraged" at regular intervals of approximately every two years for women age 75 and older whose general health and life expectancy are good (20).

Breast Cancer Survival in Older Women

There are several potential explanations for why older women experience poor breast cancer survival. These include suboptimal use of breast cancer screening, advanced stage at diagnosis, less aggressive workup, and more conservative therapy.

A series of national surveys (Behavioral Risk Factor Surveillance System, Mammography Attitudes and Usage Study, National Health Interview Survey) have documented that mammography use decreases with advancing age (11,21,22). In 1993, only 25% of women age 65 and older on Medicare had at least one mammogram (23). Rates of mammography utilization among women age 65-74, 75-84 and 85+ years were 32%, 21%, and 7%, respectively (23). Factors other than age that influence mammography use in older women include race, income, education, and state of

residence (24,25) However, having a regular provider is the most important determinant of mammography use (26-28). We examined mammography utilization among women age 65 and older and found that these sociodemographic factors remained independent predictors of mammography use even after accounting for use of primary care (29).

The stage of breast cancer at diagnosis is the most important predictor of prognosis. Women who are diagnosed while their cancer is localized to the breast experience better 5-year relative survival rate as compared with women diagnosed with more advanced disease (90% versus 64%, respectively) (13). Older women are more likely to be present with advanced disease and are more likely to go unstaged as compared with younger women disease (13,23). Furthermore, age is an independent predictor of advanced stage disease even after adjusting for other important factors (race, marital status, income, education, and source of care) (30-34).

Age has been shown to influence the diagnostic evaluation and treatment offered for breast cancer (35-38). Older women are less likely to receive diagnostic evaluations as complete or treatment as aggressive as compared with younger women. However, the poor survival experienced by the older women can primarily be attributed to their advanced stage at diagnosis since stage-specific survival is similar in all age groups and age-related treatment differences do not appear to affect survival (39).

(6) BODY

(6.1) Methods

(6.11) Data Source

We conducted a retrospective cohort study using the Linked Medicare-Tumor Registry Database (40). The linked database was jointly created by the National Cancer Institute (NCI) and the Health Care Financing Administration (HCFA) to enable researchers to conduct cancer-related health services research. The linked database contains cancer information on patients 65 years of age and older from NCI's SEER Program and Medicare enrollment and utilization information from HCFA's Medicare Statistical System. The linked database contains Medicare data from 1985 to 1994 for breast cancer cases diagnosed between 1973 and 1993.

Two Medicare utilization files are included in the linked database. First is the Medical Provider Analysis and Review (MEDPAR) file, which is a 100 percent utilization file with one record for every inpatient hospitalization or skilled nursing facility stay covered under Medicare Part A. Second is the Physicians' Claims file, which is a 100 percent utilization file with one record for every physician claim covered under Medicare Part B. Before 1991, the 100 percent Physicians' Claims file was available for only ten states. Therefore, for our study years, 1987 to 1993, data from the SEER and Medicare Programs overlap in tumor registries for three areas: Connecticut, metropolitan Atlanta, Georgia, and Seattle-Puget Sound, Washington. Specific information describing the

linkage between SEER and Medicare has been published elsewhere (40). The match rates for Connecticut, Atlanta, and Seattle were 93.3%, 94.1%, and 91.5%, respectively.

(6.12) Study Sample

Women were eligible for the study sample (n=11,399) if they received a first primary diagnosis of breast cancer between 1 January 1987 and 31 December 1993, were 67 years of age or older, and resided in Connecticut, Atlanta, or Seattle-Puget Sound. Although we selected these areas because physicians' claims were available for all cases, they also represent a geographically diverse population of older women with breast cancer. Women who were enrolled in a health maintenance organization and those with less than two full years of Medicare Part B coverage were not eligible for this study, since their physician claims data (which are required for identifying mammography use) were not available. We restricted our final study sample to women who were 67 years of age and older to ensure that all women had a full two years of Medicare utilization (claims) information before their breast cancer was diagnosed.

(6.13) Measures

We ascertained the following sociodemographic information from SEER: age at diagnosis, marital status, and SEER area. Age at diagnosis (range, 67-107 years) was categorized as 67 to 74, 75 to 84, and 85 and older. Marital status was defined as married or not at diagnosis. SEER area was classified according to the tumor registry of diagnosis: Connecticut, Atlanta, or Seattle. We used 1990 U.S. Census data to define an ecological measure of socioeconomic status: women were assigned to the median household income of their zip code of residence and grouped as $< \$25{,}000$ or $\geq \$25{,}000$.

We obtained information on race from the Medicare beneficiary enrollment file. Enrollees are classified in Medicare files as Black, White, Asian, Native American, Hispanic, or unspecified. We grouped women who were of racial/ethnic backgrounds other than Black or White together because there were too few women to permit separate analyses.

We computed a modified Charlson Comorbidity Index using Deyo's method of classifying ICD-9-CM *(International Classification of Diseases, 9th revision, Clinical Modification)* diagnosis codes from inpatient claims (41). For each woman, we identified all inpatient hospitalizations beginning two years prior to diagnosis and ending one month after diagnosis. A priori, we extended the observation period to one month past diagnosis because we expected women to have at least one hospitalization around the time of diagnosis. We classified women as 1) non-hospitalized (i.e., comorbidity could not be assessed), 2) having no comorbid conditions (Charlson Index of 0), or 3) having one or more comorbid conditions (Charlson Index ≥ 1).

We measured mammography utilization using Medicare physicians' claims. We identified all bilateral mammograms [CPT *(Physicians' Current Procedural Terminology)* procedure codes 76091 *(mammography, bilateral)* or 76092 *(screening mammography, bilateral, two films each breast)*] within two years prior to the breast cancer diagnosis. We classified women as 1) *nonusers* (n=2,029) if they had no mammograms during the entire two year period prior to diagnosis, 2) *regular users* (n=2,383) if they had at least two mammograms within the two years prior to their breast cancer diagnosis that were ten or more months apart, and 3) *peri-diagnosis users* (n=5,355) if they had their only mammogram(s) within three months before diagnosis. The peri-diagnosis users were a heterogeneous group of women whose only mammography use was close to their breast cancer diagnosis. This group includes women who had a screening mammogram, which led to their breast cancer diagnosis and those, whose mammograms were diagnostic. Therefore, analyses relating prior mammography use to breast cancer outcomes considered only nonusers and regular users, as they are clearly distinct groups.

(6.14) Logistic Regression and Hazard Ratio Modeling

Our first outcome was stage at diagnosis. We developed measures of cancer stage using both the Historical Staging System, and the TNM (tumor, node, metastases) staging system adopted by the American Joint Committee on Cancer. We utilized the latter system as the one most universally used, and providing a greater degree of differentiation across stages. The disadvantage is that we had to drop an additional 844 patients from analyses who did not have this information. We categorized late-stage disease using two classification schemes. First, women diagnosed with carcinoma in situ or Stage I tumors were classified as early-stage; those diagnosed with Stage II or greater tumors were classified as having late-stage disease. Second, we restricted late-stage disease to include only women diagnosed with Stage IIB or greater; women diagnosed with Stage IIA were reclassified as having had early-stage disease. We repeated our analyses using both classification systems and obtained similar results. We present our analyses classifying late-stage disease as Stage II or greater because they provide a more conservative estimate of the mammography-stage association.

Our second outcome was breast cancer mortality among women with invasive tumors. Women who had carcinoma in situ (n=479) were excluded from this analysis because it is unknown which tumors will progress to invasive disease. We measured survival time as the number of days from date of diagnosis until date of death or 31 December 1994 (end of follow-up). Date of death was obtained from the 1994 Medicare beneficiary enrollment file. Cause of death, obtained from SEER, captures the underlying cause listed on the death certificate. Women who had ICD-O *(International Classification of Diseases, Oncology)* codes 174.8 and 174.9 were classified as having died from breast cancer. We also calculated and present all cause mortality.

Women whose mammography use could not be categorized (788 women) or whose disease was unstaged (844 women) were excluded from the study. Overall,

there were 741 women age 67 to 74 years, 620 women 75 to 84 years, and 271 women 85 and older who met these exclusion criteria.

Follow-up for our final sample (n=9,767) ranged from one to eight years depending on the year of diagnosis. By the end of 1994, 2,332 deaths had occurred; 889 deaths were attributed to breast cancer (385 women 67 to 74 years, 390 women 75 to 84 years, and 114 women 85 years and older).

(6.15) Simulated Case-Control Analysis

After review of the recent literature on uses of observational data to address issues of benefit in screening, our research team has developed the following methodology innovative in its application in the field of mammography screening. It combines the recently applied analytic methods of propensity score matching and instrumental variable analysis to simulate a case-control study and adjust for selection bias in our administrative dataset cohort (42,43).

Randomized trials are viewed as the "gold standard" in research. However, this is one of the cases where these types of trials cannot be conducted due to ethical constraints. Treatment and control groups potentially differ in their baseline characteristics, which makes the results susceptible to bias. Only by adjusting for these differences in baseline characteristics can we determine the true treatment effect.

The standard method for analysis of a dichotomous outcome is a logistic regression. The exposure of interest is included as a predictor of outcome, with other covariates included to control for baseline differences. The odds ratio is then estimated, and tested to determine if it is statistically different from 1, which indicates no effect. This method will work if either all relevant baseline risk factors have been measured or the unmeasured factors are highly correlated with those that are measured, and if the relationship between risk factors and outcomes are correctly specified in the model. However, it is almost never reasonable to believe that all the important risk factors have been measured and it is often difficult to know whether relationships between measured risk factors and the outcomes have been correctly specified.

One method of adjustment that handles possible model mis-specification is the use of propensity score matching. This process involves calculating for each case a propensity score, which indicates the likelihood (or "propensity") that each case is a user of screening mammography. Then, a sample of the user and non-user groups is taken matching on the propensity score. As shown by Rosenbaum and Rubin (44), matching on the propensity score will result in a sub-sample of the study and control group that are well balanced in terms of observed risk factors. The standard analysis approach is then applied to the sub-sample data to determine the effect of using screening mammography on stage at diagnosis, after adjusting for any residual differences in observed covariates that remain after matching. The third method that we employ is an instrumental variable approach. This approach is commonplace in

econometrics, dating back to the 1920s, and is becoming increasingly popular in the health services outcomes research field as well. For a variable, or group of variables, to be considered a good instrument, it should be neither associated with the outcome beyond its effect on exposure, nor with unmeasured confounders, after adjusting for those factors already in the model.

(6.16) Propensity Score and Instrumental Variable Model Development

For the standard analysis, we developed a logistic regression model to predict stage at diagnosis from user status, controlling for age, race, comorbidity, regional income, and primary care visits. The c-statistic was determined as a measure of predictive validity.

For the propensity score approach, a logistic regression model including age, race, comorbidity, regional income, region, and primary care visits was used to determine the propensity of being a user. A split into deciles based on the propensity scores was done. To develop the matched samples, within each decile a random sample of the larger group (regular users or non-users) was taken in order to get the same number as was in the smaller group. The reduced samples were then combined into the analytic dataset.

A second model excluding region was also calculated to determine if the specification of propensity differed by region. The results did not differ so this alternate method of generating propensity scores was not pursued further. To examine the extent to which the above matching resulted in samples of regular users and non-users more comparable in terms of baseline characteristics, p-values for chi-squared tests of independence were calculated for the categorical risk factors variables. We calculated p-values for independent sample t-tests for equivalence of population means for numeric risk factors.

For the instrumental variable analysis, the first step is to determine which variable or variables are to be used as instruments. In our situation, a candidate for an instrument must be a predictor of whether someone is a regular user of mammography with no residual predictive power on stage at diagnosis, after controlling for the other covariates in the model, including the propensity of being a regular user. Angrist, Imbens, and Rubin (45) describe how this can be broken down into two necessary conditions. We consider these conditions in the context of our example, using region as our instrument. The first necessary condition is that the effect of moving from one region to another, conditional on user status and the measured covariates, does not change the expected outcome. For example, we would expect that a woman with certain characteristics (primary care visits, age, race, etc.) receiving regular screening in Seattle, would have the same likelihood of early stage disease diagnosed from mammography had she moved to Atlanta or Connecticut. If this assumption were not met, it would imply that, after conditioning on observed covariates, follow-up after a positive mammogram in one region is different than follow-up in another region. There

is no evidence to suggest this is the case. The second condition is that there is random, or at least ignorable, assignment to whether or not the woman was a regular user. Failure to meet this assumption would imply that there are unmeasured variables that differ across regions that are associated with race, age, primary care visits (which might be thought of as a proxy for access), income, or comorbidities.

These covariates are also associated with whether or not someone becomes a regular user. Though there is no way to empirically validate this last assumption, it seems reasonable in the context of our example. The instrumental variable approach involves a two-stage model. In the first stage, covariates plus the instruments are used to predict user status. The predicted probability of being a regular user is then used in lieu of user status as an independent variable in the second stage, along with any measured covariates. Variables used as instruments in the first stage model are excluded, since these variables are assumed to effect the outcome only through their association with user status. The coefficient associated with predicted user is a measure of the impact of use.

As a measure the strength of the instrumental variables, we used results from the first stage model. The odds ratio as well as the Wald chi-squared value was used. Staiger and Stock (46) suggest that a partial F-statistic of about 10 is sufficient to not be a weak instrument. Instead we use Wald chi-squared statistics from our logistic regression models, calculated as the difference in the -2 log likelihoods between the full and reduced (excluding the instrument) models to predict user status. The difference between the partial F (square of a t-distribution) and Wald chi-squared (square of a normal distribution) should be insignificant due to our sample size, so that the same rule of thumb may apply. In addition, the odds ratio of each group is presented as another measure of the strength of the association.

(6.17) Literature Review

In order to compare our odds ratios and hazards ratios to the data available from the existing randomized clinical trials, we conducted a literature review of all reports in the previous 15 years from both randomized clinical trials, and from case control and cohort studies which addressed mammography use. We reviewed this literature for age specific comparison, which would allow us to compare the odds ratios estimates for women 65 and older with women 50-65 in the clinical trials. We hope to plot these results along with our own for women 67-74 years, 75-84 years, and 85 years and greater. We planned to plot our results with both the adjusted hazards ratio, as well as the adjusted hazards ratio allowing 1.25 years for lead time bias.

Our review of the literature revealed only one report published which stratified the 50 year and older age group into smaller age groupings and found equivalent benefits in 50-60 and 60 and older women. However, with no specific age group overlapping the youngest women in our population, we are unable to complete this analysis to date. We

have been unsuccessful in obtaining more age-specific data to conduct further comparisons with our results.

(6.2) Results

Characteristics of the study sample (n=9,767) are presented by age group at diagnosis in Table 1. Overall, 47% of women were aged 67 to 74 years at the time of their breast cancer diagnosis, 42% were 75 to 84 years, and 11% were 85 years or older.

Overall, 21% of women had no mammograms within two years prior to their breast cancer diagnosis (nonusers), 24% of women had at least two mammograms within two years preceding diagnosis that were ten or more months apart (regular users), and 55% had their only mammogram(s) within three months prior to their diagnosis (peri-diagnosis users). Figure 1 presents the percentage of women who were nonusers and regular users of mammography according to age at diagnosis. The proportion of women who were peri-diagnosis users was similar across the age groups and is not displayed. Regular mammography use decreased with advancing age at diagnosis such that the women in the oldest age group were substantially less likely to undergo regular mammograms: 29% of women 67 to 74 years, 23% of women 75 to 84 years, and 10% of women 85 years or older were regular users. Although in the two youngest age groups, the proportion of nonusers was similar (18% of women 67 to 74 years and 21% of women 75 to 84 years) and less than the proportion of regular users, the reverse was true for the oldest women. One-third of women 85 years and older did not undergo mammography within two years before their diagnosis.

Figure 2 presents the distribution of stage at diagnosis according to age at diagnosis. Within each age group, most women presented with Stage I or Stage II breast cancers. The distribution of disease among women in the two younger groups is nearly identical, except that fewer women 75 to 84 years presented with carcinoma in situ compared with women 67 to 74 years (8% versus 12%, respectively). However, there is a noticeable shift in the distribution of disease among the oldest women, characterized by the greater frequency of Stage II and Stage III cancers diagnosed in women 85 years and older. Late-stage (i.e., Stage II or greater) breast cancer was diagnosed in 41% and 45% of women 67 to 74 years and 75 to 84 years, respectively, and in 53% of women 85 years or older,

Figure 3 presents the percentage of nonusers and regular users of mammography who were diagnosed with late-stage disease for each age group. Within each age group, nonusers were significantly more likely to be diagnosed with late-stage disease than regular users. Furthermore, the proportion of nonusers who were diagnosed with late-stage disease increased with advancing age (49% aged 67 to 74 years, 60% aged 75 to 84 years, and 69% aged 85 years or older). In contrast, the proportion of regular users who presented with late-stage disease at diagnosis was substantially lower (28%) and was similar across age groups.

Table 2 presents the odds ratios for late-stage disease comparing nonusers with regular users of mammography for all women and for each age group separately. These analyses were performed to determine whether the relation between prior mammography use and stage at diagnosis is significant for older women of different age groups. Prior mammography use was strongly associated with stage at diagnosis for all women and women in each age group. Even after adjusting for factors that have been found to be associated with late-stage disease at diagnosis, including age at diagnosis, race, marital status, income of zip code of residence, and comorbid conditions, lack of mammography use remained a significant predictor of late-stage at diagnosis in all women (adjusted OR, 3.12 [95% CI, 2.74-3.58)] and within each age group: 67 to 74 years (adjusted OR, 2.46 [95% CI, 2.04-2.98]); 75 to 84 years (adjusted OR, 3.64 [95% CI, 2.96-4.48]); and 85 years or older (adjusted OR, 6.87 [95% CI, 3.97-11.90]).

Table 3 presents overall 5-year survival estimates (i.e., deaths from all causes) following diagnosis by stage at diagnosis and age group. Survival decreased with later stage at diagnosis. Furthermore, survival decreased steadily with advancing age within each cancer stage. Table 4 presents the hazard ratios for breast cancer mortality, comparing nonusers with regular users of mammography for all women and for each age group separately. Table 4 also presents results demonstrating the potential effect of a lead time of 1.25 years. The results in Table 4 focus on breast cancer mortality because one would expect that mammography would primarily impact on deaths from breast cancer. After adjusting for sociodemographic factors and comorbidity, nonusers were at significantly greater risk of death from breast cancer than regular users and had greater risk of dying from breast cancer within each age group. Consideration of lead time somewhat diminished the magnitude of the hazard ratio, but nonusers of mammography continued to be at increased risk of dying from breast cancer. Our findings remained significant for all women and for the two youngest age groups. Although the point estimate remained increased for the oldest women, it no longer achieved statistical significance.

(6.21) Results of Propensity Score and Instrumental Variable Analysis

The final sample size was 4667 women. Of the 4667 women, 11 were excluded from analyses due to missing zip codes, which made the zip code-income matching impossible. Of the remaining 4656, 1354 were diagnosed with late stage cancer and 3302 with early stage cancer; 2516 were regular users and 2140 were non-users.

The standard model controlling for age at diagnosis, race, comorbidities, regional income, region, and number of primary care had a c-statistic of 0.68. The conclusion, based on this model, is that regular users have 2.97 times the odds of being diagnosed at early stage relative to non-users (95% CI: 2.56, 3.45). These results are comparable to the logistic regression model reported above (6.2) using the entire N= 9767 sample.

In the propensity score analysis, the model to predict user status using age, race, comorbidities, regional income, region, and primary care visits as independent variables had a c-statistic of 0.78. Table 5 shows the pre-and post-sampling number of cases by decile and compares baseline characteristics of regular users and non-users.

We can see that the matching resulted in much more balance between the groups in terms of the measured covariates. Each of these variables showed a statistical difference based on user status before sampling yet no difference after sampling. The exception to this is stage at diagnosis, which was not included in the propensity score model, since it occurs after use status and hence could not influence the likelihood of being a regular user.

When the logistic regression was run on the matched sample, the odds ratio associated with user status was 3.27 (95% CI: 2.72, 3.93), a result similar to the pre-sampling results from the standard logistic regression analysis.

The result from the instrumental variable approach using region as an instrument produced an odds ratio of 3.01 (95% CI: 1.09, 8.34), again similar to the standard analysis. The chi-squared statistic for not using region as a predictor of user was 53.0 ($p<.001$). The odds ratio for predicting user status from region is 1.34 for Seattle and 0.64 for Atlanta, both using Connecticut as the reference group.

In addition, we considered two other instruments. First was race (black vs. other). The odds ratio for early detection for regular users relative to non-users is 3.69 (95% CI: 0.88, 15.57). The chi-squared statistic for not using race as a predictor of user was 5.74 ($p=.017$). The odds ratio for predicting user status from race is 0.67, indicating that blacks have two-thirds the odds of non-blacks to become a regular user of mammography.

The second alternative instrumental variable analysis we conducted was using primary care visits as the instrument. The resulting odds ratio is 2.45 (95%CI: 1.80, 3,35). The chi-squared statistic for not using primary care visits as a predictor of user was 824 ($p<.001$). The odds ratio for predicting user status from number of primary care visits is 3.87, 9.73, and 14.70 for 1-3, 4-12, and more than 12 primary care visits, respectively, all relative to those women with no primary care visits.

(6.3) Discussion

In the absence of data on older women from randomized controlled trials, we sought to improve our understanding of the relationship between prior mammography use and breast cancer outcomes in older women. We found that women with breast cancer aged 67 years and older were significantly more likely to be diagnosed with Stage II or greater disease if they were nonusers of mammography than if they were

regular users. Women who were nonusers of mammography were also at greater risk of dying from their breast cancer than those who were regular users. Most importantly, these findings persist with advancing age at diagnosis.

We have shown that regular mammography use is associated with earlier diagnosis of breast cancer in older women. In fact, within each age group studied, we found that in Stage I cancers were more common among regular users and that the proportion of women who presented with Stage II or greater disease was substantially lower for regular users than for nonusers.

Even though reduced breast cancer mortality is the ultimate goal of breast cancer screening, some have argued that intermediate measures, such as stage at diagnosis, are useful for evaluating the utility of screening (47). Our data demonstrate a significant reduction in Stage II or greater tumors among women who were regular mammography users. It is well established that stage at diagnosis is the most important predictor of survival and that stage is inversely correlated with survival. Therefore, these results suggest that regular users should have a more favorable prognosis because they are diagnosed earlier in the disease process. Previously, the Swedish Two-Country Trial demonstrated that a 25% reduction in advanced staged breast cancers for screened women translated to a 30% reduction in breast cancer mortality (48).

Although Medicare claims data have been used effectively to measure mammography (23,29,49), there are some potential limitations. First, our study is limited to women enrolled in fee-for-service settings, as Medicare data do not capture services rendered to HMO enrollees. Few women were enrolled in managed care during our study years, however, the proportion of Medicare HMO enrollees increased from 4% to 13% between 1990 and 1997 (40, 50).

Second, Medicare reimbursement policies have changed over time. Medicare began reimbursing providers for biennial screening mammography in 1991 and annual screening mammography in 1998. Although Medicare only paid for diagnostic mammograms prior to 1991, studies show that providers were performing screening mammograms and billing Medicare under the diagnostic procedure code both before and after the change in reimbursement (23, 29, 49). Nevertheless, we cannot determine whether an individual mammogram was done for screening or diagnostic purposes. To address this issue, we examined how women used mammography over time. We defined our measure of mammography use to identify two distinct groups: 1) women who had no evidence of mammography use during the two years before diagnosis, and 2) those who demonstrated a pattern of regular use by having had at least two mammograms that were at least 10 months apart. We selected 10 months as a clinically reasonable interval to assume that women were undergoing screening and were not being followed for a suspicious lump. Although women with a pattern of regular use seemed to be using mammography for screening, we do not know which women had their cancer detected by symptoms and confirmed with diagnostic mammography.

Due to the observational nature of this study, two potential sources for bias inherent to evaluations of cancer screening must considered. These sources are lead-time bias, and bias due to differences in baseline characteristics between users and nonusers of mammography. We conducted a number of innovative analyses to assess for these biases.

Lead-time bias artificially extends the survival time of screened women by advancing the date of diagnosis and is the primary methodological limitation to using post-diagnosis survival as an outcome of cancer screening. Since we do not know any individual's lead time or whom had their tumors diagnosed clinically or through screening, we sought to explore the potential effect of lead-time bias on our survival results by allowing for a lead-time of 1.25 years for each regular user (51). We found that adjustment for lead-time diminished the magnitude of the hazard ratio, but that nonusers of mammography continued to be at increased risk of dying from breast cancer. Our findings remained significant for all women and for the two youngest age groups. However, for the oldest women, the point estimate for the hazard ratio remained increased, but the confidence limits included unity—possibly due to the fewer women in this age group.

Baseline differences between the two groups is another major potential source of bias when analyzing observational data. We first conducted a standard logistic regression model to adjust for known confounders, including age, stage, race and comorbidities. The standard analysis gives us results to which we can compare the innovative analyses in order to determine the amount of model mis-specification (when results are compared to the propensity score results) and the effect of unmeasured confounders (when the results are compared to the instrumental variable analysis). It is important to note that the odds ratio of 2.97 from the standard analysis is a measure of effect of use for women of a type similar to those in the entire sample.

The propensity score matching method produced an odds ratio of 3.27. While this effect is slightly larger than that from the standard analysis, the difference is clearly not statistically significant, based on the confidence intervals. The propensity score matching approach is an estimate of the impact of being a regular user of mammography for the sample that is similar in terms of the measured covariates that were included in the propensity generation model. Since this result is so close to that of the standard model, it implies that there is not a large bias that has come into effect in the standard model due to a biased selection of women who have self-selected to be regular users or non-users of mammography.

The main instrument under consideration was region. The significance of region on predicting user status was quite strong (Chi-squared of 53, $p<.001$ from the reduced model), indicating that we do not have a weak instrument. For otherwise identical women with average characteristics, the effect of being in Seattle rather than Connecticut increases the odds of mammogram use by 34%. This is consistent with

shown geographic variation in practice patterns, in particular those of breast cancer care (52, 53). A weak instrument results in a bias in estimating program impact toward the estimate from the standard approach. If our instrument were weak, we could not conclude that similarity of instrumental variable estimates and standard analysis estimates implied that there were not unmeasured confounders.

The instrumental variable analysis, using region as the instrument, produced an odds ratio of 3.01. Again, this result is similar to that of the standard analysis, implying that there is no large bias coming from unmeasured variables. This value measures the impact among those whose behavior changes depending on their region, with all else being equal. The confidence interval for this value (1.09, 8.34) is much wider than that of the standard model (2.56, 3.45). This is expected, because the instrumental variable approach depends on much less variation in user status, i.e., exclusively on differences in user status caused by regional differences, than do the alternative models. With less variation in user status, the estimate of the standard error will be larger.

In summary, all three analyses, the standard regression, the propensity score matching, and the instrumental variable analysis using region as the instrument, produced very similar results. We conclude that the standard regression analysis is valid. There is little model mis-specification, either from measured variables, as seen via the propensity score matching, or from unmeasured variables, as seen via the instrumental variable analysis.

In summary, this study describes the relationship between prior mammography use, cancer stage at diagnosis, and breast cancer mortality. Our data suggest that women who fail to undergo mammography are more often diagnosed with Stage II or greater breast cancers and are at increased risk of dying from breast cancer. Moreover, breast cancer was an important cause of death; breast cancer was the cause in 38% of all deaths in our study sample. These data support the use of regular mammography in older women up to age 85 years, and suggest that mammography can reduce breast cancer mortality for older women, even for women age 85 and older.

(6.4) Statement of Work

We have met all objectives of the project.

(7) KEY RESEARCH ACCOMPLISHMENTS

1. Oral presentation to the Era of Hope Meeting

2. Manuscript accepted for publication: McCarthy EP, Burns RB, Freund KM, Ash AS, Shwartz M, Marwill SL, Moskowitz MA. Mammography use, breast cancer stage at diagnosis, and survival among older women. J Am Geriat Soc 2000;48:1226-1233.

3. Developed innovative methodology using propensity scores, and instrumental variables.

4. Manuscript submitted to Health Services Outcomes Research.

5. Mr. Posner received a student paper award from the Health Policy and Statistics section of the American Statistical Association.

6. Mr. Posner received first place award at the Boston University Graduate Student Science Day.

7. Oral presentation at the Joint Statistical Meetings in Atlanta in August 2001.

(8) REPORTABLE OUTCOMES

Manuscripts

1. McCarthy EP, Burns RB, Freund KM, Ash AS, Shwartz M, Marwill SL, Moskowitz MA. Mammography use, breast cancer stage at diagnosis, and survival among older women. J Am Geriat Soc 2000;48:1226-1233.

2. Posner MA, Ash AS, Shwartz M. Freund KM, Moskowitz MA.Comparing standard regression, propensity score matching, and instrumental variables methods for determining the effectiveness of mammography in older women. Health Services Outcomes (submitted)

Abstracts

1. McCarthy EP, Burns RB, M Freund KM, Marwill SL, Ash AS, Shwartz M, Moskowitz MA. Does regular mammography use improve breast cancer outcomes in older women? J Gen Intern Med 1997;12(Suppl 1):76.

2. McCarthy EP, Freund KM, Burns RB, Ash AS, Shwartz M, Moskowitz M. Is mammography useful in older women? Proceedings of the Era of Hope, Department of Defense Breast Cancer Research Program Meeting, June 8-11, 2000, Volume I:178.

(9) CONCLUSIONS

In the absence of randomized clinical trial data, observational data is the only available data to address the issue of whether to continue mammography screening in women over age 70 years. Our observational study has utilized a large, prospectively collected database in order to assess the association between regular mammography use and stage and mortality from breast cancer. The two major limitations of

observational data are, first, confounding and bias due to differences between women who receive regular mammography and those who do not. The second major limitation of observational data is the lead time bias, which can result in the appearance of a benefit to screening when none exists.

We have used innovative methods to address these two major limitations of observational data. We have conducted traditional logistic regression and hazard ratio modeling. In addition we performed a propensity score analysis and instrumental analysis to adjust for potential unmeasured confounding. We have developed methods to try and estimate lead time bias, and adjust for this bias. All adjustments for potential confounding consistently find that women between the ages of 67 and 85 have a lower risk of late stage disease and a lower risk of breast cancer death with regular mammography use. We therefore conclude that our data support the recommendation of continued mammography screening in elderly women.

(10) REFERENCES:

1. Yancik R, Ries LG, Yates JW. Breast cancer and aging women: A population-based study of contrast in stage, surgery and survival. Cancer 1989;63:976-981.

2. Santariano WA, Bell SH, Swanson GM. The severity of breast cancer at diagnosis: A comparison of age and extent of disease in black and white women. Am J Public Health 1986;76:770-782.

3. Homes FF, Hearne E. Cancer stage-to-end relationship: Implications for cancer screening in the elderly. J Am Geriatr Soc 1981;29:55-57.

4. Goodwill JS, Same JM, Key CR, Humble C, Kutvirt D, Hunt C. Stage at diagnosis of cancer varies with the age of the patient. J Am Geriatr Soc 1986;34;20-26.

5. Ries LAG, Miller BA, Hankey BF, Kosary CL, Harras A, Edwards BK, eds. *SEER Cancer Statistics Review, 1973-1991. Tables and Graphs.* Bethesda, MD: National Cancer Institute, 1994, NIH Publication 94-2789.

6. Faulk RM, Sickles EA, Sollitto RA, Ominsky SH, Galvin HB, Frankel SD. Clinical efficacy of mammographic screening in the elderly. Radiol 1995;194:193-197.

7. Satariano WA, Belle SH, Swanson GM. The severity of breast cancer at diagnosis: A comparison of age and extent of disease in black and white women. Am J Public Health 1986;76:779-82.

8. Clark GM, Osborne CK, McGuire WL. Correlations between estrogen receptor, progesterone receptor, and patient characteristics in human breast cancer. J Clin Oncol 1984;2:1102-1109.

9. Clark GM. The biology of breast cancer in older women. J Gerontol 1992;47:19-23.

10. EPESE, 1982-83. U.S. DATA: National Center for Health Statistics: Vital Statistics of the United States, 1982, Vol. II, Mortality, Part A. DHHS Publication No. (PHS)86-1122. Public Health Service, Washington, DC: U.S. Government Printing Office, 1986.

11. Kopans DB. Screening mammography in women over age 65. J Gerontol 1992;47:59-62.

12. Kerlikowske K, Grady D, Barclay J, Sickles EA, Eaton A, Ernester V. Positive predictive value of screening mammography by age and family history of breast cancer. JAMA 1993;270;2444-2450.

13. Dodd GD. American Cancer Society guidelines on screening for breast cancer: An overview. Cancer 1992;69:1885-1887.

14. Branch LG, Guralnik JM, Foley DJ, et al. Active life expectancy for 10,000 men and women in three communities. J Gerontol 1991;4M145-150.

15. Mandelblatt JS, Wheat ME, Monane M, et al. Breast cancer screening for elderly women with and without comorbid conditions. Ann Intern Med 1992;116:722-730.

16. Mushlin AI, Fintor L. Is screening for breast cancer cost-effective? Cancer 1991;69:1957-1962.

17. Brown ML. Economic considerations in breast cancer screening of older women. J Genontol 1992;47:51-58.

18. Council on Scientific Affairs. Early detection of breast cancer. JAMA 1984;252:3008-3011.

19. Woolf SH, Kamerow DB, Lawrence RS, Medalie JH, Estes EH. The Periodic health examination of older adults: The recommendation of the U.S. Preventive Services Health Task Force. J Am Geriatr Soc 1990;38:933-942.

20. Screening recommendations of the Forum Panel. J Gerontol 1992;47(special issue):5.

21. Centers for Disease Control. Use of mammography – United States, 1900. MMWR 1990;39:621-630.

22. Centers for Disease Control. Ackermann SP, Brackbill RM, Bewerse BA, Sanderson LM. Cancer screening behaviors among U.S. Women: Breast cancer, 1987-1989, and cervical cancer, 1988-1989. MMWR 1992;41:17-25.

23. Centers for Disease Control. Trontell AE. Use of mammography services by women aged ≥ 65 years enrolled in Medicare – United States, 1991-19993. MMWR 1995;44(41)777-781.

24. Coleman EA, Feuer EJ, and The NCI Breast Cancer Screening Consortium. Ann Intern Med 1992;117:961-966.

25. Mor V, Pacala JT, Rakowski W. Mammography for older women: Who uses, who benefits? J Gerontol 1992;47:43-49.

26. The NCI Breast Cancer Screening Consortium. Screening mammography: A missed clinical opportunity? Results of the NCI Breast Cancer Screening Consortium and National Health Interview Studies. JAMA 1990;264:54-58.

27. Fox SA, Murata PJ, Stein JA. The effect of physician compliance on screening mammography for older women. Arch Inter Med 1991;151:50-56.

28. Zapka JG, Stoddard AM, Costanza ME, Greene HL. Breast cancer screening by mammography: Utilization and associated factors. Am J Public Health 1989;79:1499-1502.

29. Burns RB, McCarthy EP, Freund KM, Marwill SL, Shwartz M, Ash A, Moskowitz MA. Black women receive less mammography even with similar use of primary care. Ann Intern Med 1996;125:171:-182.

30. Mandelblatt J, Andrews H, Kerner J, Zauber A, Burnett W. Determinants of late stage diagnosis of breast and cervical cancer: The impact of age, race, social class, and hospital type. Am J Public Health 1991;81:646-649.

31. Farley TA, Flannery JT. Late-stage diagnosis of breast cancer in women of lower socioeconomic status: Public health implications. Am J Public Health 1989;79:1508-1512.

32. Wells BL, Horm JW. Stage at diagnosis in breast cancer: Race and socioeconomic factors. Am J Public Health 1992;82:1383-1385.

33. Kulkarni PR, Orth D, Vernon SW, Jackson GL, Lairson D, Davis DR. Stage at diagnosis of breast cancer. Comparison in a fee-for-service and health maintenance organization practice. Med Care 1989;27:608-622.

34. Riley GF, Potosky AL, Lubitz JD, Brown ML. Stage of cancer at diagnosis for Medicare HMO and fee-for-service enrollees. Am J Public Health 1994;84:1598-1604.

35. Chu J, Diehr P, Feigl P, et al. The effect of age on the care of women with breast cancer in community hospitals. J Gerontol 1987;42:185-190.

36. Greenfield S, Blanco DM, Elashoff RM, Ganz PA. Patterns of care related to age of breast cancer patients. JAMA 1987;257:2766-2770.

37. Samet J, Hunt WC, Key C, Humble CG, Goodwin JS. Choice of cancer therapy varies with age of patient. JAMA 1986;255:3385-3390.

38. Silliman RA, Guadagnoli E, Weitberg AG, Mor V. Age as a predictor of diagnostic and initial treatment intensity in newly diagnosed breast cancer patients. J Gerontol 1989;33:M46-50.

39. Bergman L, Kluck HM, van Leeuwen MA, et al. The influence of age on treatment choice and survival of elderly breast cancer patients in South-eastern Netherlands: A population-based study. Eur J Cancer 1992;28A:1475-1480.

40. Potosky AL, Riley GF, Lubitz JD, Mentnech RM, Kessler LG. Potential for cancer related health services research using a Linked Medicare-Tumor Registry Database. Med Care 1993;31:732-48.

41. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative database. J Clin Epidemiol 1992;45:613-9.

42. Trials and Case-Control Studies in Evaluating the Effectiveness of Screening Mammography. J Clin Epidemiol 1998;51:81-91.

43. Shwartz M, Ash AS, Iezzoni, LI. The use of non-experimental Studies to evaluate surgical procedures and other interventions: the challenge of risk adjustment. In Souba WW, Wilmore DW, eds. Surgical Research, San Diego: Academic Press, 2001.

44. Rosenbaum PR, Rubin DB. The Central Role of the Propensity Score in Observational Studies for Causal Effects. Biometrika 1983;70:41-55.

45. Angrist JD, Imbens, GW, Rubin, DB. Identification of Causal Effects Using Instrumental Variables. JASA 1996;96:444-72.

46. Sheiner LB, Rubin DB. Intention-to-Treat Analysis and the Goals of Clinical Trials. Clin Pharm Therap 1995;57:6-15.

47. Day NE. Quantitative approach to the evaluation of screening programs. World J Surg. 1989;12:3-8.

48. Tabar L, Gad A, Holmberg L, Ljungquist U. Significant reduction in advanced breast cancer. Results of the first seven years of mammography screening in Kopparberg, Sweden. Diagn Imag Clin Med. 1985;54:I58-164.

49. Blustein J. Medicare coverage, supplemental insurance, and the use ofmammography by older women. N Engl J Med 1995;332:1138-43.

50. Bureau of Data Management and Strategy, Health Care Financing Administration. 1997 Data Compendium. Baltimore, MD: U.S. Department of Health and Human Services, Health Care Financing Administration; 1997. Publication no. 03395.

51. Tabar L. Fagerberg G, Chen HH, et al. Efficacy of breast cancer screening by age. New results from the Swedish Two-County Trial. Cancer 1995;75:2507-17.

52. Nattinger AB, Gottlieb MS, Veum J, Yagnke DL, Goodwin JS. Geographic variation in the use of breast conserving treatment for breast cancer. N Engl J Med. 1992;326:1102-1107.

53. Nattinger AB, Gottlieb MS. Hoffman RG, Walker AP, Goodwin JS. Minimal increase in use of breast-conserving surgery from 1986 to 1990. Med Care. 1996;24:479-489.

(11) LIST OF PERSONNEL FUNDED ON PROJECT

Karen Freund MD MPH
Mark Moskowitz MD
Arlene Ash, PhD
Michael Posner MS
Marianne Prout MD MPH
Anita Harrington MA
M Geraldine Rankin
Patricia Lewis

Table 1. Characteristics of the Study Sample by Age at Diagnosis

	Age at Diagnosis			
	67 to 74	75 to 84	≥ 85	Total
	(n=4,609)	(n=4,072)	(n=1,086)	(n=9,767)
	n (%)	n (%)	n (%)	n (%)
SEER Area*				
Connecticut	2110 (46)	1992 (49)	536 (50)	4638 (48)
Seattle	1687 (37)	1392 (34)	384 (35)	3463 (35)
Atlanta	812 (17)	688 (17)	166 (15)	1666 (17)
Race				
White	4236 (92)	3785 (93)	1020 (94)	9041 (93)
Black	208 (4)	178 (4)	38 (3)	424 (4)
Other	165 (4)	109 (3)	28 (3)	302 (3)
Married at Diagnosis†				
No	2251 (49)	2803 (69)	975 (90)	6029 (62)
Yes	2358 (51)	1269 (31)	111 (10)	3738 (38)
Median Income of Zip Code				
≥ $25,000	4158 (91)	3678 (91)	976 (90)	8812 (91)
< $25,000	432 (9)	386 (9)	107 (10)	925 (9)

	Age at Diagnosis			
	67 to 74 (n=4,609)	75 to 84 (n=4,072)	≥ 85 (n=1,086)	Total (n=9,767)
	n (%)	n (%)	n (%)	n (%)
Comorbidity[†]				
Non-hospitalized	1174 (25)	1001 (25)	298 (27)	2473 (25)
0	2559 (56)	2126 (52)	460 (42)	5145 (53)
1	627 (14)	640 (16)	224 (21)	1491 (15)
≥ 2	249 (5)	305 (7)	104 (10)	658 (7)

[*]$p = 0.018$.

[†]$p < 0.001$.

Figure 1. Prior Mammography Use By Age at Diagnosis

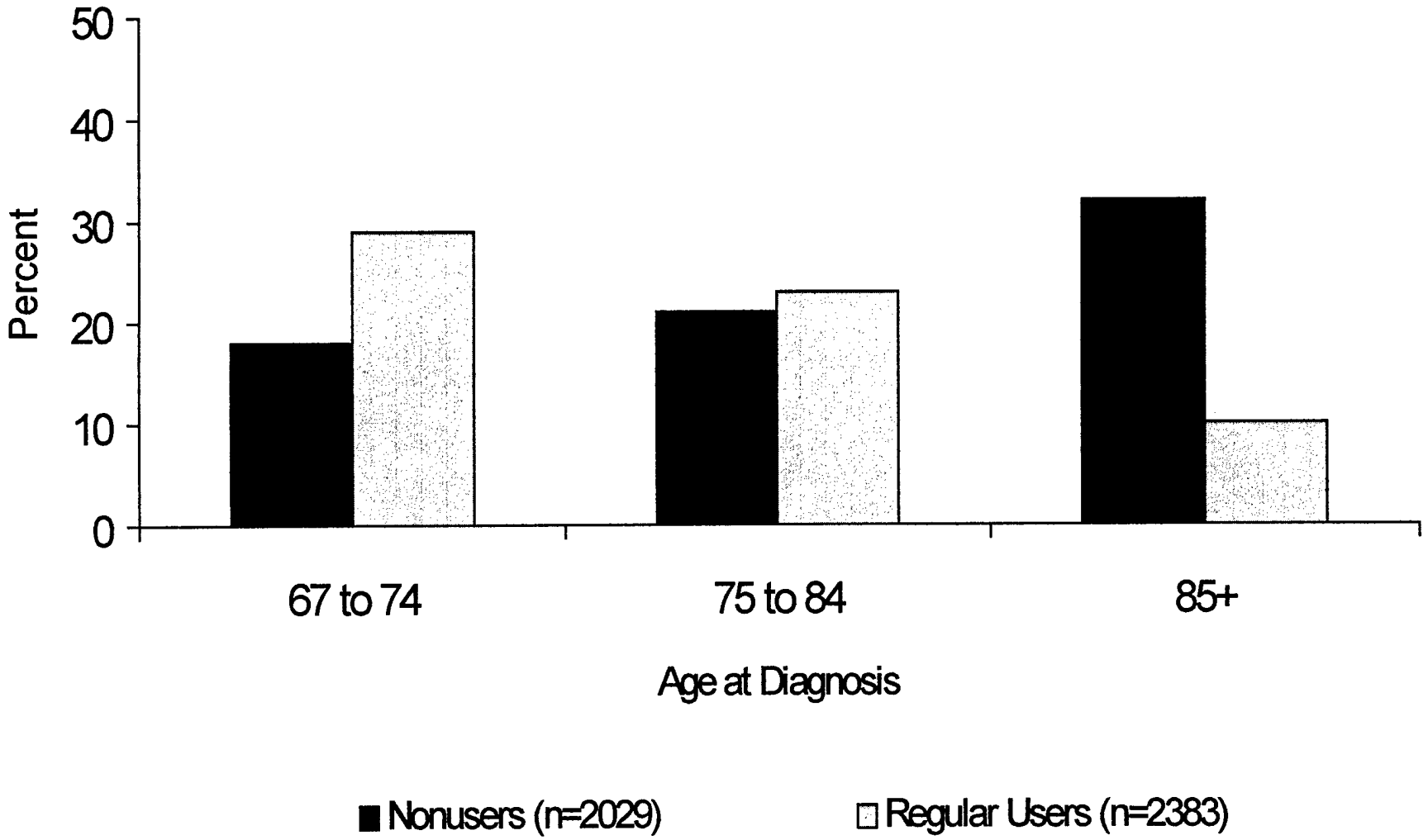

Figure 2. Stage at Diagnosis By Age at Diagnosis

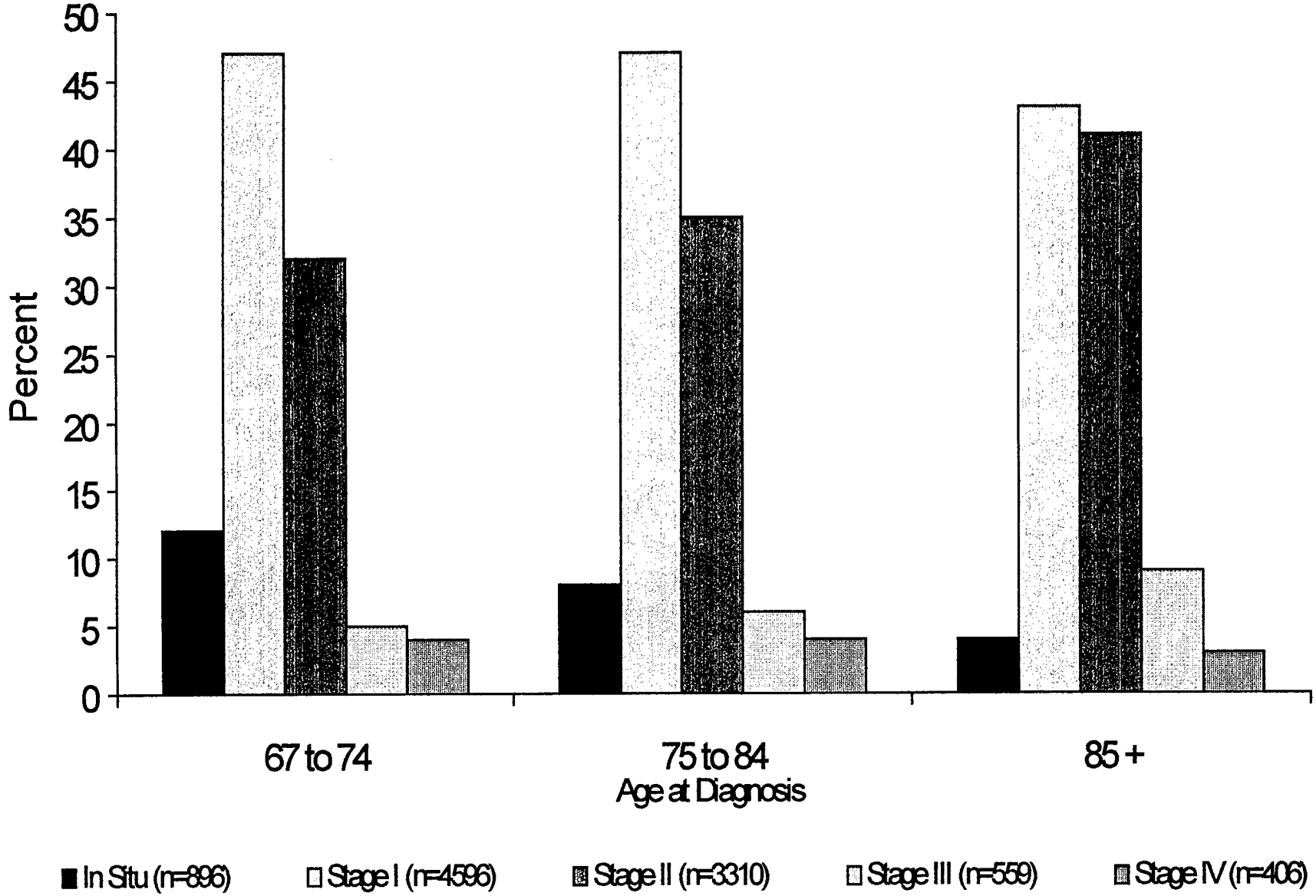

Figure 3. Percentage of Women with Stage ≥ II Disease By Prior Mammography Use and Age at Diagnosis

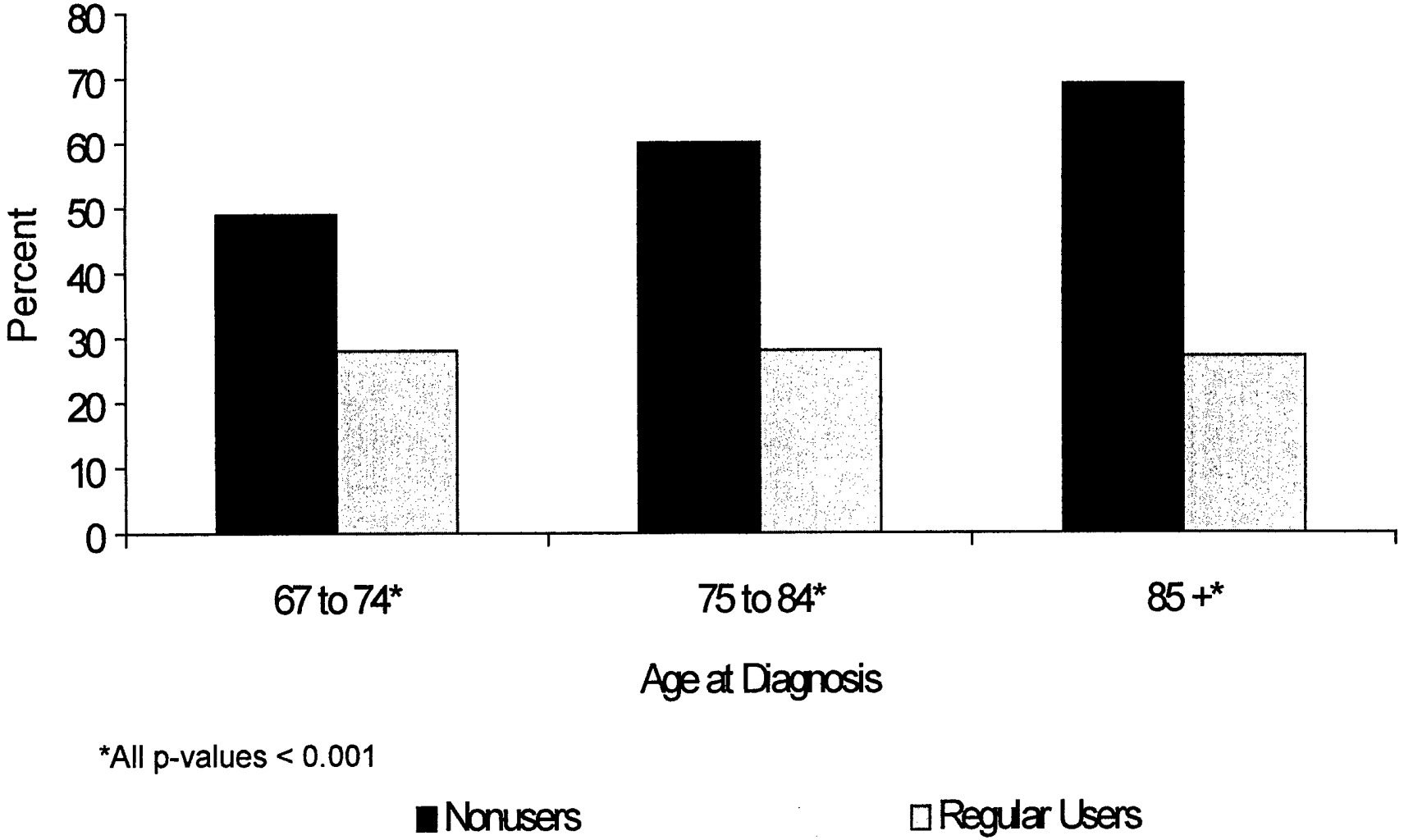

Table 2. Crude and Adjusted Odds Ratios for Late Stage Disease

Nonusers Compared with Regular Users (n = 4,412)

	Stage ≥ II at Diagnosis	
	Crude	**Adjusted***
	OR (95% CI)	**OR (95% CI)**
All Women (n=4412)	3.36 (2.96-3.81)	3.12 (2.74-3.58)
Age 67 to 74 (n=2167)	2.43 (2.03-2.92)	2.46 (2.04-2.98)
Age 75 to 84 (n=1790)	3.74 (3.07-4.55)	3.64 (2.96-4.48)
Age ≥ 85 (n=455)	6.25 (3.86-10.12)	6.87 (3.97-11.90)

*Adjusted for age at diagnosis as a continuous variable, race, marital status, income of ZIP Code, and comorbidity.

Table 3. Relation of Stage at Diagnosis to Five-Year Survival Estimates By Age at Diagnosis Among Nonusers and Regular Users Combined (n = 3,933)*

Stage at Diagnosis and Age	n	5-year Estimated Survival (SE)	Log Rank P-value†
Stage I			
67 to 74	1083	0.877 (0.013)	
75 to 84	856	0.842 (0.016)	
≥ 85	168	0.496 (0.052)	0.0001
Stage II			
67 to 74	567	0.785 (0.021)	
75 to 84	535	0.620 (0.026)	
≥ 85	185	0.345 (0.043)	0.0001
Stage III/IV			
67 to 74	217	0.362 (0.040)	
75 to 84	238	0.286 (0.036)	
≥ 85	84	0.225 (0.055)	0.039

* Women with carcinoma in situ were excluded from these analyses.

†Log Rank tests differences in overall survival by age at diagnosis.

Table 4. Crude and Adjusted Risk of Breast Cancer Mortality by Age at Diagnosis

Nonusers Compared with Regular Users (n = 3,933)*

	Breast Cancer Mortality			
	From the Date of Diagnosis		Assuming a Lead Time of 1.25 Years for Regular Users	
	Crude	Adjusted†	Crude	Adjusted†
	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)
All Women (n=3933)	3.53 (2.80-4.45)	3.38 (2.65-4.32)	2.36 (1.87-2.98)	2.28 (1.79-2.91)
Age 67 to 74 (n=1867)	3.18 (2.27-4.46)	2.94 (2.08-4.17)	2.14 (1.52-3.01)	2.14 (1.51-3.02)
Age 75 to 84 (n=1627)	3.69 (2.58-5.27)	3.95 (2.72-5.74)	2.41 (1.68-3.46)	2.47 (1.70-3.58)
Age ≥ 85 (n=437)	2.71 (1.22-6.04)	2.29 (1.00-5.26)	1.74 (0.78-3.90)	1.45 (0.63-3.32)

* Women with carcinoma in situ were excluded from these analyses.

†Adjusted for age at diagnosis as a continuous variable, race, marital status, income of ZIP Code, comorbidity, and year of diagnosis. Proportional Hazards models were stratified on SEER area.

‡Hazard ratio (95% confidence interval).

Table 5: Propensity Score Matching Results						
	Pre-Sampling			Post-Sampling		
	Non-User	User	p-value	Non-User	User	p-value
Total Sample	2140	2516		1274	1274	
Decile 1	416	**57**		57	**57**	
Decile 2	339	**89**		89	**89**	
Decile 3	359	**136**		136	**136**	
Decile 4	239	**205**		205	**205**	
Decile 5	**204**	305		**204**	204	
Decile 6	**158**	287		**158**	158	
Decile 7	**135**	321		**135**	135	
Decile 8	**112**	366		**112**	112	
Decile 9	**100**	379		**100**	100	
Decile 10	**78**	371		**78**	78	
Age at Diagnosis	77.2	74.5	0.001	75.8	75.6	0.428
Age at Diagnosis						
67-69	37.8%	62.2%		48.9%	51.1%	
70-74	39.1%	60.9%		51.4%	48.6%	
75-79	44.1%	55.9%		49.7%	50.3%	
80-84	51.6%	48.4%		49.1%	50.9%	
85+	76.5%	23.5%	0.001	50.0%	50.0%	0.919
Charlson Comorbidities						
Not Hospitalized	43.6%	56.4%		48.7%	51.3%	
Hosp, No Comorbidities	43.0%	57.0%		50.3%	49.7%	
At Least One Comorbidity	56.9%	43.1%	0.001	51.0%	49.0%	0.699
Race						
Black	63.2%	36.8%		55.9%	44.1%	
Non-Black	45.1%	54.9%	0.001	49.7%	50.3%	0.143
Income (Median of Zip Code)	$42,030	$41,137	0.061	$42,004	$41,727	0.672
Regional Income						
Top 40%	44.1%	55.9%		50.4%	49.6%	
Lower 60%	47.3%	52.7%	0.036	49.7%	50.3%	0.748
Primary Care Visits	4.9	10.5	0.001	7.6	8.2	0.074
Primary Care Visits						
None	83.1%	16.9%		51.2%	48.8%	
1-3	55.7%	44.3%		50.0%	50.0%	
4-12	32.8%	67.2%		50.1%	49.9%	
13+	26.5%	73.5%	0.001	49.2%	50.8%	0.952
Location						
Seattle	35.8%	64.2%		49.6%	50.4%	
Atlanta	55.0%	45.0%		49.5%	50.5%	
Connecticut	50.0%	50.0%	0.001	50.5%	49.5%	0.908
Stage						
Early (TNM 0 or I)	38.2%	61.8%		41.6%	58.4%	
Late (TNM II, III, or IV)	64.8%	35.2%	0.001	69.6%	30.4%	0.001

Mammography Use, Breast Cancer Stage at Diagnosis, and Survival Among Older Women

Ellen P. McCarthy, PhD, MPH,[*] Risa B. Burns, MD, MPH,[*] Karen M. Freund, MD, MPH,[†] Arlene S. Ash, PhD,[†] Michael Shwartz, PhD,[‡] Sandra L. Marwill, MD, MPH,[§] and Mark A. Moskowitz, MD[†]

BACKGROUND: Women age 65 years and older account for most newly diagnosed breast cancers and deaths from breast cancer. Yet, older women are least likely to undergo mammography, perhaps because mammography's value is less well demonstrated in older women.

OBJECTIVE: To investigate the relationship between prior mammography use, cancer stage at diagnosis, and breast cancer mortality among older women with breast cancer.

DESIGN: Retrospective cohort study using the Linked Medicare-Tumor Registry Database.

SETTING: Population-based data from three geographic areas included in the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) program.

PARTICIPANTS: Women aged 67 and older diagnosed with a first primary breast cancer, from 1987 to 1993, residing in Connecticut, metropolitan Atlanta, Georgia, or Seattle-Puget Sound, Washington.

MEASUREMENTS: Medicare claims were reviewed and women were classified according to their mammography use during the 2 years before diagnosis: nonusers (no prior mammograms), regular users (at least two mammograms at least 10 months apart), or peri-diagnosis users (only mammogram(s) within 3 months before diagnosis). Mammography utilization was linked with SEER data to determine stage at diagnosis and cause of death. Our main outcome variables were (1) stage at diagnosis, classified as early (in situ/Stage I) or late (Stage II or greater), and (2) breast cancer mortality, measured from diagnosis until death from breast cancer or end of the follow-up period (December 31, 1994).

RESULTS: Older women who were nonusers of mammography were diagnosed with breast cancer at Stage II or greater more often than regular users (adjusted odds ratio (OR), 3.12; 95% confidence interval (CI), 2.74–3.58). This association was present within each age group studied. Nonusers of mammography were at significantly greater risk of dying from their breast cancer than regular users for all women (adjusted hazard ratio (HR), 3.38; 95% CI, 2.65–4.32) and for women within each age group. Even assuming a lead time of 1.25 years, nonusers of mammography continued to be at increased risk of dying from breast cancer. Our findings remained significant for all women and for the two youngest age groups (67–74 years, 75–85 years), although the benefit was no longer statistically significant for the oldest women (85 years and older).

CONCLUSIONS: Older women who undergo regular mammography are diagnosed with an earlier stage of disease and are less likely to die from their disease. These data support the use of regular mammography in older women and suggest that mammography can reduce breast cancer mortality in older women, even for women age 85 and older. J Am Geriatr Soc 48:1226–1233, 2000.

Key words: mammography; breast cancer; stage at diagnosis; older adults; Medicare

Breast cancer is an important public health problem in the United States, particularly among older women (age 65 years and older). Each year, more than 180,000 American women develop breast cancer and more than 40,000 die from the disease.[1] Older women bear the greatest burden of disease because the risks of developing and dying from breast cancer rise sharply with advancing age.[2] The National Cancer Institute's (NCI) Surveillance, Epidemiology, and End Results (SEER) Program reports that older women account for 48% of newly diagnosed invasive breast cancers and 58% of breast cancer deaths.[3] The SEER Program also documents a 12% increase in the breast cancer mortality rate in older women from 1973 to 1994, despite a decline in overall rate.[3] This may be partially due to the fact that older women are more likely to present with advanced breast cancer at diagnosis,[2,4–6] possibly because they are less likely to undergo regular mammography.[7]

From the [*]Division of General Medicine and Primary Care, Department of Medicine, Harvard Medical School, Beth Israel Deaconess Medical Center, Boston, Massachusetts; [†]Section of General Internal Medicine, Evans Department of Medicine, Boston Medical Center, Boston, Massachusetts; [‡]Operations Management, Boston University School of Management, Boston, Massachusetts; [§]Gillette Medical Department, One Gillette Park, South Boston, Massachusetts.
Presented at the Society of General Internal Medicine annual meeting, May 1997.

Grant Support: The United States Army Medical Research and Materiel Command Breast Cancer Program Grant No. DMAD-17-97-1-7121. The views expressed are solely those of the authors.
Address correspondence and reprint requests to Ellen P. McCarthy, PhD, Division of General Medicine and Primary Care, Department of Medicine, Beth Israel Deaconess Medical Center, 330 Brookline Avenue, Libby-326, Boston, MA 02215.

0002-8614/00/$3.50

Early detection with mammography has been shown to reduce breast cancer mortality by 20–39% in women aged 50–69 years.[8–12] Despite this striking reduction in mortality, it is unknown whether mammography continues to be useful beyond age 70 or at what age, if any, breast cancer screening and early detection is no longer of value.[13] Unfortunately, none of the randomized controlled mammography trials included women over 74 years of age. Although one mammography trial (the Swedish Two-County Trial) included women up to age 74 years, there was inadequate power to establish a mortality reduction specifically for women over age 69 years.[12] Therefore, breast cancer screening guidelines for older women are based on extrapolating data from women under age 70 and mammography recommendations vary.

Without clear scientific evidence, older women and their physicians are left to make decisions on an individual basis. Because it is unlikely that a randomized controlled trial will ever be implemented in women over age 70, conclusions about the value of mammography in older women will most likely be inferred from observational studies, such as the one reported here. Therefore, we sought to improve our understanding of the relationship between previous mammography use and (1) cancer stage at diagnosis, and (2) breast cancer mortality among older women diagnosed with breast cancer in NCI's SEER Program.

METHODS

Data Source

We conducted a retrospective cohort study using the Linked Medicare-Tumor Registry Database.[14] The linked database was jointly created by the NCI and the Health Care Financing Administration (HCFA) to enable researchers to conduct cancer-related health services research. The linked database contains cancer information on patients 65 years of age and older from NCI's SEER Program and Medicare enrollment and utilization information from HCFA's Medicare Statistical System. The linked database contains Medicare data from 1985 to 1994 for breast cancer cases diagnosed between 1973 and 1993.

Two Medicare utilization files are included in the linked database. First is the Medical Provider Analysis and Review (MEDPAR) file, which is a 100% utilization file with one record for every inpatient hospitalization or skilled nursing facility stay covered under Medicare Part A. Second is the Physicians' Claims file, which is a 100% utilization file with one record for every physician claim covered under Medicare Part B. Before 1991, the 100% Physicians' Claims file was available for only 10 states. Therefore, for our study years, 1987–1993, data from the SEER and Medicare programs overlap in tumor registries for three areas: Connecticut, metropolitan Atlanta, Georgia, and Seattle-Puget Sound, Washington. Specific information describing the linkage between SEER and Medicare has been published elsewhere.[14] The match rates for Connecticut, Atlanta, and Seattle were 93.3%, 94.1%, and 91.5%, respectively.

Study Sample

Women were eligible for the study sample (n = 11,399) if they received a first primary diagnosis of breast cancer between 1 January 1987 and 31 December 1993, were 67 years of age or older, and resided in Connecticut, Atlanta, or Seattle-Puget Sound. Although we selected these areas because physicians' claims were available for all cases, they also represent a geographically diverse population of older women with breast cancer. Women who were enrolled in a health maintenance organization and those with less than 2 full years of Medicare Part B coverage were not eligible for this study, because their physician claims data (which are required for identifying mammography use) were not available. We restricted our final study sample to women who were 67 years of age and older to ensure that all women had a full 2 years of Medicare utilization (claims) information before their breast cancer was diagnosed.

Measures

We ascertained the following sociodemographic information from SEER: age at diagnosis, marital status, and SEER area. Age at diagnosis (range, 67–107 years) was categorized as 67–74, 75–84, and 85 and older. Marital status was defined as married or not at diagnosis. SEER area was classified according to the tumor registry of diagnosis: Connecticut, Atlanta, or Seattle. We used 1990 US Census data to define an ecological measure of socioeconomic status: women were assigned to the median household income of their zip code of residence and grouped as <$25,000 or ≥$25,000.

We obtained information on race from the Medicare beneficiary enrollment file. Enrollees are classified in Medicare files as black, white, Asian, Native American, Hispanic, or unspecified. We grouped women who were of racial/ethnic backgrounds other than black or white together because there were too few women to permit separate analyses.

We computed a modified Charlson Comorbidity Index using Deyo's method of classifying ICD-9-CM *(International Classification of Diseases, 9th revision, Clinical Modification)* diagnosis codes from inpatient claims.[15] For each woman, we identified all inpatient hospitalizations beginning 2 years before diagnosis and ending 1 month after diagnosis. A priori, we extended the observation period to 1 month past diagnosis because we expected women to have at least one hospitalization around the time of diagnosis. We classified women as (1) nonhospitalized (i.e., comorbidity could not be assessed), (2) having no comorbid conditions (Charlson Index = 0), or (3) having one or more comorbid conditions (Charlson Index ≥ 1).

We measured mammography utilization using Medicare physicians' claims. We identified all bilateral mammograms [CPT (Physicians' Current Procedural Terminology) procedure codes 76091 (mammography, bilateral) or 76092 (screening mammography, bilateral, two films each breast)] within 2 years before the breast cancer diagnosis. We classified women as (1) nonusers (n = 2029) if they had no mammograms during the entire 2-year period before diagnosis, (2) regular users (n = 2383) if they had at least two mammograms within the 2 years before their breast cancer diagnosis that were 10 or more months apart, and (3) peri-diagnosis users (n = 5355) if they had their only mammogram(s) within 3 months before diagnosis. The peri-diagnosis users were a heterogeneous group of women whose only mammography use was close to their breast cancer diagnosis. This group includes women who had a screening mammogram that led to their breast cancer diagnosis and those whose mammograms were diagnostic. Therefore, analyses relating prior mammography use to breast cancer outcomes

considered only nonusers and regular users, as they are clearly distinct groups.

Our first outcome was stage at diagnosis. We measured cancer stage using the TNM (tumor, node, metastases) staging system adopted by the American Joint Committee on Cancer. We categorized late-stage disease using two classification schemes. First, women diagnosed with carcinoma in situ or Stage I tumors were classified as early-stage; those diagnosed with Stage II or greater tumors were classified as having late-stage disease. Second, we restricted late-stage disease to include only women diagnosed with Stage IIB or greater; women diagnosed with Stage IIA were reclassified as having had early-stage disease. We repeated our analyses using both classification systems and obtained similar results. We present our analyses classifying late-stage disease as Stage II or greater because they provide a more conservative estimate of the mammography-stage association.

Our second outcome was breast cancer mortality among women with invasive tumors. Women who had carcinoma in situ (n = 479) were excluded from this analysis because it is unknown which tumors will progress to invasive disease. We measured survival time as the number of days from date of diagnosis until date of death or December 31, 1994 (end of follow-up). Date of death was obtained from Medicare beneficiary enrollment files. Cause of death, obtained from SEER, captures the underlying cause listed on the death certificate. Women who had ICD-O (International Classification of Diseases, Oncology) codes 174.8 and 174.9 were classified as having died from breast cancer.

Women whose mammography use could not be categorized (788 women) or whose disease was unstaged (844 women) were excluded from the study. Overall, there were 741 women age 67–74 years, 620 women 75–84 years, and 271 women 85 and older who met these exclusion criteria.

Follow-up for our final sample (n = 9767) ranged from 1 to 8 years depending on the year of diagnosis. By the end of 1994, 2332 deaths had occurred; 889 deaths were attributed to breast cancer (385 women 67–74 years, 390 women 75–84 years, and 114 women 85 years and older).

Statistical Analysis

All statistical analyses were performed using SAS statistical software version 6.12.[16] We performed each analysis once for all women and again for women within each age group. We compared women across age groups with respect to sociodemographic factors, comorbidity, stage at diagnosis, and prior mammography use. Chi-square statistics and *t* tests were used to identify characteristics at diagnosis that varied significantly with age at diagnosis.

Multivariable logistic regression was used to estimate crude and adjusted odds of late-stage disease for women who failed to undergo mammography compared with women who used regular mammography.[17] The odds ratio (OR) for prior mammography use and the corresponding 95% confidence intervals (CI) were estimated from the β coefficient and standard error from the logistic models.[17] Multivariable logistic models adjusted for factors previously found to be related to stage at diagnosis including age at diagnosis, race, marital status, income of zip code of residence, and comorbid conditions.[18] For models fit to each age group, we adjusted for age at diagnosis as a continuous variable to account for any residual confounding with age.

To better understand overall survival (i.e., death from all causes) in our study sample, we computed Kaplan-Meier estimates of 5-year survival by age group for each stage at diagnosis. We combined women with Stage III and IV disease to have sufficient numbers for meaningful analysis. Because this analysis describes survival regardless of cause of death, only women who were alive at the end of follow-up were censored. The log-rank test was used to identify differences in overall survival by age group within each stage stratum.[19]

To further examine the relationship of mammography use and survival, we hypothesize that mammography use should primarily affect breast cancer-related deaths. Therefore, we fit stratified Cox proportional hazards regression models to estimate the crude and adjusted risk of death from breast cancer for women who failed to undergo mammography compared with women who used mammography regularly. All models were stratified by SEER area to account for any lack of proportionality among the three tumor registries by allowing the underlying hazard to differ. In these analyses, women were also censored when they died from causes other than their breast cancer. Each hazard ratio (HR) (i.e., relative risk of mortality) for prior mammography use and its corresponding 95% CI was estimated from the β coefficient and standard error from a Cox model.[19]

Analyses of postdiagnosis survival in relation to cancer screening are subject to lead time bias in which a woman whose disease is diagnosed earlier through screening will live longer "following diagnosis" simply due to earlier detection. Unfortunately, we do not know any individual's lead time or which women had tumors diagnosed clinically or through screening. We explored the potential effect of lead time bias on our survival results by estimating the risk of dying from breast cancer for nonusers compared with regular users after allowing for a lead time of 1.25 years for each regular user. The number 1.25 years is approximately one-half the mean sojourn time (i.e., on average the maximum lead time achievable) for women 70–74 years in the Swedish Two-County Trial.[20]

RESULTS

Characteristics of the study sample (n = 9767) are presented by age group at diagnosis in Table 1. Overall, 47% of women were aged 67–74 years at the time of their breast cancer diagnosis, 42% were 75–84 years, and 11% were 85 years or older.

Overall, 21% of women had no mammograms within 2 years before their breast cancer diagnosis (nonusers), 24% of women had at least two mammograms within 2 years preceding diagnosis that were 10 or more months apart (regular users), and 55% had their only mammogram(s) within 3 months before their diagnosis (peri-diagnosis users). Figure 1 presents the percentage of women who were nonusers and regular users of mammography according to age at diagnosis. The proportion of women who were peri-diagnosis users was similar across the age groups and is not displayed. Regular mammography use decreased with advancing age at diagnosis such that the women in the oldest age group were substantially less likely to undergo regular mammograms: 29% of women 67–74 years, 23% of women 75–84 years, and 10% of women 85 years or older were regular users. Although in the two youngest age groups, the proportion of nonusers was similar (18% of women 67–74 years and 21% of women 75–84 years) and less than the proportion of regular users,

Table 1. Characteristics of the Study Sample by Age at Diagnosis

	Age at Diagnosis			
	67 to 74 (n = 4609)	75 to 84 (n = 4072)	≥85 (n = 1086)	Total (n = 9767)
	n	(%)	*n*	(%)
SEER area*				
Connecticut	2110 (46)	1992 (49)	536 (50)	4638 (48)
Seattle	1687 (37)	1392 (34)	384 (35)	3463 (35)
Atlanta	812 (17)	688 (17)	166 (15)	1666 (17)
Race				
White	4236 (92)	3785 (93)	1020 (94)	9041 (93)
Black	208 (4)	178 (4)	38 (3)	424 (4)
Other	165 (4)	109 (3)	28 (3)	302 (3)
Married at diagnosis†				
No	2251 (49)	2803 (69)	975 (90)	6029 (62)
Yes	2358 (51)	1269 (31)	111 (10)	3738 (38)
Median income of zip code				
≥$25,000	4158 (91)	3678 (91)	976 (90)	8812 (91)
<$25,000	432 (9)	386 (9)	107 (10)	925 (9)
Comorbidity†				
Nonhospitalized	1174 (25)	1001 (25)	298 (27)	2473 (25)
0	2559 (56)	2126 (52)	460 (42)	5145 (53)
1	627 (14)	640 (16)	224 (21)	1491 (15)
≥2	249 (5)	305 (7)	104 (10)	658 (7)

*$P = .018$.
†$P < .001$.

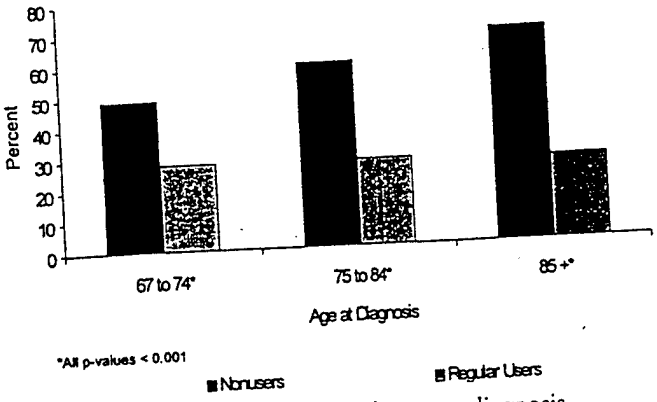

Figure 1. Prior mammography use by age at diagnosis.

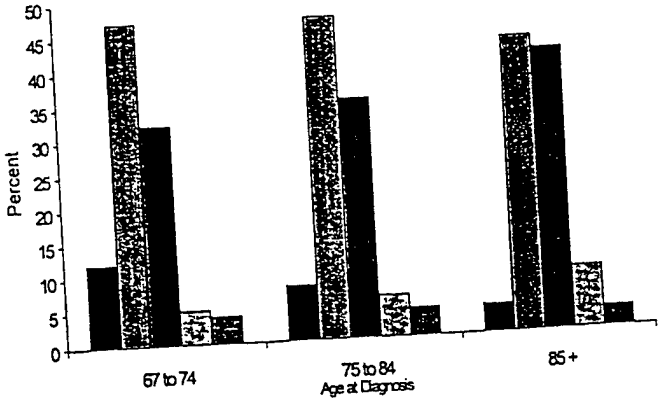

Figure 2. Stage at diagnosis by age at diagnosis.

the reverse was true for the oldest women. One-third of women 85 years and older did not undergo mammography within 2 years before their diagnosis.

Figure 2 presents the distribution of stage at diagnosis according to age at diagnosis. Within each age group, most women presented with Stage I or Stage II breast cancers. The distribution of disease among women in the two younger groups is nearly identical, except that fewer women 75–84 years presented with carcinoma in situ compared with women 67–74 years (8% versus 12%, respectively). However, there is a noticeable shift in the distribution of disease among the oldest women, characterized by the greater frequency of Stage II and Stage III cancers diagnosed in women 85 years and older. Late-stage (i.e., Stage II or greater) breast cancer was diagnosed in 41% and 45% of women 67–74 years and 75–84 years, respectively, and in 53% of women 85 years or older.

Figure 3 presents the percentage of nonusers and regular users of mammography who were diagnosed with late-stage disease for each age group. Within each age group, nonusers were significantly more likely to be diagnosed with late-stage disease than regular users. Furthermore, the proportion of nonusers who were diagnosed with late-stage disease increased with advancing age (49% aged 67–74 years, 60% aged 75–84 years, and 69% aged 85 years or older). In contrast, the proportion of regular users who presented with late-stage disease at diagnosis was substantially lower (28%) and was similar across age groups.

Table 2 presents the odds ratios for late-stage disease comparing nonusers with regular users of mammography for

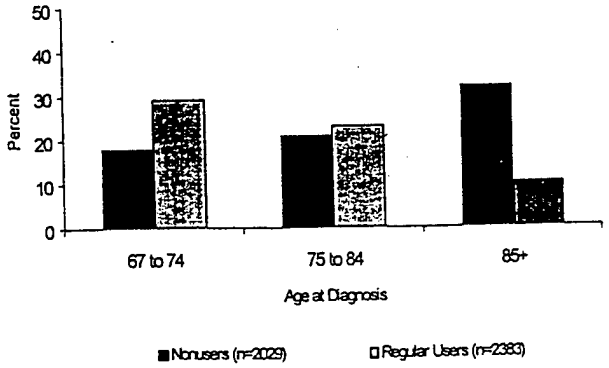

Figure 3. Percentage of women with Stage II or greater disease by prior mammography use and age at diagnosis.

Table 2. Crude and Adjusted Odds Ratios for Late Stage Disease: Nonusers Compared with Regular Users (n = 4,412)

	Stage ≥II at Diagnosis	
	Crude OR (95% CI)	Adjusted* OR (95% CI)
All Women (n = 4412)	3.36 (2.96–3.81)	3.12 (2.74–3.58)
Age 67–74 (n = 2167)	2.43 (2.03–2.92)	2.46 (2.04–2.98)
Age 75–84 (n = 1790)	3.74 (3.07–4.55)	3.64 (2.96–4.48)
Age ≥85 (n = 455)	6.25 (3.86–10.12)	6.87 (3.97–11.90)

*Adjusted for age at diagnosis as a continuous variable, race, marital status, income of ZIP Code, and comorbidity.

all women and for each age group separately. These analyses were performed to determine whether the relation between prior mammography use and stage at diagnosis is significant for older women of different age groups. Prior mammography use was strongly associated with stage at diagnosis for all women and women in each age group. Even after adjusting for factors that have been found to be associated with late-stage disease at diagnosis, including age at diagnosis, race, marital status, income of zip code of residence, and comorbid conditions, lack of mammography use remained a significant predictor of late-stage at diagnosis in all women (adjusted OR, 3.12; 95% CI, 2.74–3.58) and within each age group: 67–74 years (adjusted OR, 2.46; 95% CI, 2.04–2.98); 75–84 years (adjusted OR, 3.64; 95% CI, 2.96–4.48); and 85 years or older (adjusted OR, 6.87; 95% CI, 3.97–11.90).

Table 3 presents overall 5-year survival estimates (i.e., deaths from all causes) following diagnosis by stage at diagnosis and age group. Survival decreased with later stage at diagnosis. Furthermore, survival decreased steadily with advancing age within each cancer stage.

Table 4 presents the hazard ratios for breast cancer mortality, comparing nonusers with regular users of mammography for all women and for each age group separately. Table 4 also presents results demonstrating the potential effect of a lead time of 1.25 years. The results in Table 4 focus on breast cancer mortality because one would expect that mammography would primarily impact on deaths from breast cancer. After adjusting for sociodemographic factors and comorbidity, nonusers were at significantly greater risk of death from breast cancer than regular users and had greater risk of dying from breast cancer within each age

Table 3. Relation of Stage at Diagnosis to Five-Year Survival Estimates by Age at Diagnosis Among Nonusers and Regular Users Combined (n = 3933)*

Stage at Diagnosis and Age	n	5-year Estimated Survival (SE)	Log-Rank *P*-value†
Stage I			
67–74	1083	0.877 (0.013)	
75–84	856	0.842 (0.016)	
≥85	168	0.496 (0.052)	.0001
Stage II			
67–74	567	0.785 (0.021)	
75–84	535	0.620 (0.026)	
≥85	185	0.345 (0.043)	.0001
Stage III/IV			
67–74	217	0.362 (0.040)	
75–84	238	0.286 (0.036)	
≥85	84	0.225 (0.055)	.039

*Women with carcinoma in situ were excluded from these analyses.
†Log-rank tests differences in overall survival by age at diagnosis.

group. Consideration of lead time somewhat diminished the magnitude of the hazard ratio, but nonusers of mammography continued to be at increased risk of dying from breast cancer. Our findings remained significant for all women and for the two youngest age groups. Although the point estimate remained increased for the oldest women, it no longer achieved statistical significance.

DISCUSSION

In the absence of data on older women from randomized controlled trials, we sought to improve our understanding of the relationship between prior mammography use and breast cancer outcomes in older women. We found that women with breast cancer aged 67 years and older were significantly more likely to be diagnosed with Stage II or greater disease if they were nonusers of mammography than if they were regular users. Women who were nonusers of mammography were also at greater risk of dying from their breast cancer than those who were regular users. Most importantly, these findings persist with advancing age at diagnosis.

We have shown that regular mammography use is associated with earlier diagnosis of breast cancer in older women. In fact, within each age group studied, we found that in situ and Stage I cancers were more common among regular users and that the proportion of women who presented with Stage II or greater disease was substantially lower for regular users than for nonusers. Two studies have demonstrated similar results. A case control study of older women conducted in the United States suggested an association between screening mammography and a reduction in metastatic breast cancer, but lacked sufficient power to demonstrate significance.[21] Faulk and coworkers examined the clinical efficacy of mammography among women aged 65 years and older compared with women aged 50–64 years and found that mammography was at least as effective in detecting breast cancers with a favorable prognosis in older women on several measures. They found that mammography detected slightly smaller tumors in older women, that were more often axillary node negative, and in an earlier stage.[22]

Table 4. Crude and Adjusted Risk of Breast Cancer Mortality by Age at Diagnosis: Nonusers Compared with Regular Users (n = 3933)*

	Breast Cancer Mortality			
	From the Date of Diagnosis		Assuming a Lead Time of 1.25 Years for Regular Users	
	Crude HR‡ (95% CI)	Adjusted† HR (95% CI)	Crude HR (95% CI)	Adjusted† HR (95% CI)
All Women (n = 3933)	3.53 (2.80–4.45)	3.38 (2.65–4.32)	2.36 (1.87–2.98)	2.28 (1.79–2.91)
Age 67–74 (n = 1867)	3.18 (2.27–4.46)	2.94 (2.08–4.17)	2.14 (1.52–3.01)	2.14 (1.51–3.02)
Age 75–84 (n = 1627)	3.69 (2.58–5.27)	3.95 (2.72–5.74)	2.41 (1.68–3.46)	2.47 (1.70–3.58)
Age ≥85 (n = 437)	2.71 (1.22–6.04)	2.29 (1.00–5.26)	1.74 (0.78–3.90)	1.45 (0.63–3.32)

*Women with carcinoma in situ were excluded from these analyses.
†Adjusted for age at diagnosis as a continuous variable, race, marital status, income of ZIP Code, comorbidity, and year of diagnosis. Proportional Hazards models were stratified on SEER area.
‡Hazard ratio (95% confidence interval).

Even though reduced breast cancer mortality is the ultimate goal of breast cancer screening, some have argued that intermediate measures, such as stage at diagnosis, are useful for evaluating the utility of screening.[23] Our data demonstrate a significant reduction in Stage II or greater tumors among women who were regular mammography users. It is well established that stage at diagnosis is the most important predictor of survival and that stage is inversely correlated with survival. Therefore, these results suggest that regular users should have a more favorable prognosis because they are diagnosed earlier in the disease process. Previously, the Swedish Two-Country Trial demonstrated that a 25% reduction in advanced staged breast cancers for screened women translated to a 30% reduction in breast cancer mortality.[24]

Indeed we found a consistent lower risk of death from breast cancer among regular users of mammography overall and within each age group. Few other studies have examined the relationship between mammography utilization and breast cancer mortality for older women, and most have relied on observational data. The only randomized study to offer insight into screening for women aged 70–74 years is the Swedish Two-County trial, which showed a persistent 34% reduction in breast cancer mortality for women 50–74 years after 13 years of follow-up.[20] Although poor compliance precluded age-specific mortality analyses for women over age 70,[1] more than half of the women over age 70 who died of breast cancer were among those who had refused screening.[25] A follow-up study of women in the Breast Cancer Detection Demonstration Project found that the observed number of breast cancer deaths among women aged 60–74 years was 26% less than the expected based on national data.[26] Case control studies in the Netherlands have also evaluated the efficacy of screening mammography in older women by comparing a population-based screening program in Nijmegen to a neighboring city without a formal screening program. Although initial analyses suggested a modest effect of mammography,[27,28] a recent analysis estimated that regular mammography reduced breast cancer mortality in older women by approximately 45%.[29]

Although Medicare claims data have been used effectively to measure mammography,[7,18,30–32] there are some potential limitations. First, our study is limited to women enrolled in fee-for-service settings, as Medicare data do not capture services rendered to HMO enrollees. Few women were enrolled in managed care during our study years, however, the proportion of Medicare HMO enrollees increased from 4% to 13% between 1990 and 1997.[14,33]

Second, Medicare reimbursement policies have changed over time. Medicare began reimbursing providers for biennial screening mammography in 1991 and annual screening mammography in 1998. Although Medicare only paid for diagnostic mammograms before 1991, studies show that providers were performing screening mammograms and billing Medicare under the diagnostic procedure code both before and after the change in reimbursement.[7,30–32] Nevertheless, we cannot determine whether an individual mammogram was done for screening or diagnostic purposes. To address this issue, we examined how women used mammography over time. We defined our measure of mammography use to identify two distinct groups: (1) women who had no evidence of mammography use during the 2 years before diagnosis, and (2) those who demonstrated a pattern of regular use by having had at least two mammograms that were at least 10 months apart. We selected 10 months as a clinically reasonable interval to assume that women were undergoing screening and were not being followed for a suspicious lump. Although women with a pattern of regular use seemed to be using mammography for screening, we do not know which women had their cancer detected by symptoms and confirmed with diagnostic mammography.

Due to the observational nature of this study, three potential sources for bias inherent to evaluations of cancer screening must considered. These sources of bias are overdetection, length time, and lead time.[34] Overdetection bias occurs when screening detects potentially clinically insignificant tumors. To address this issue, we approached our survival analyses conservatively by excluding women diagnosed with carcinoma in situ to minimize the potential affect of overdetection bias.

Length time bias occurs when the underlying tumor growth rate differs between screened and unscreened groups such that slow-growing, less aggressive tumors are more likely to be detected with screening and fast-growing, aggressive tumors are more likely to be detected clinically. For this reason, length time bias is of particular concern in studies that include prevalent cancers found on initial screening examina-

tions. Length time bias is less of a concern in our study for two reasons. First, our sample consisted of women with incident breast cancer. Second, older women generally have less aggressive tumors than younger women.[35] We have only limited information on tumor characteristics. However, regular users and nonusers seem similar with respect to pathologic grade; only 18% of regular users and 20% of nonusers had poor or undifferentiated tumors.

As mentioned previously, lead time bias artificially extends the survival time of screened women by advancing the date of diagnosis and is the primary methodologic limitation to using postdiagnosis survival as an outcome of cancer screening. Because we do not know any individual's lead time or who had their tumors diagnosed clinically or through screening, we sought to explore the potential effect of lead time bias on our survival results by allowing for a lead time of 1.25 years for each regular user.[20] We found that adjustment for lead time diminished the magnitude of the hazard ratio, but that nonusers of mammography continued to be at increased risk of dying from breast cancer. Our findings remained significant for all women and for the two youngest age groups. However, for the oldest women, the point estimate for the hazard ratio remained increased, but the confidence limits included unity—possibly due to the fewer women in this age group.

Finally, it is important to note the following limitations when interpreting our breast cancer mortality results. First, follow-up for our study sample ranged from 1 to 8 years. It is unknown whether the mortality difference that we observed between nonusers and regular users would persist with additional years of follow-up. Second, although we considered the effect of lead time using an estimate of 1.25 years based on data for women aged 70–74 years,[20] it is possible that lead time may be longer. Unfortunately, we do not have enough follow-up to examine longer lead time estimates.

In summary, this study describes the relationship between prior mammography use, cancer stage at diagnosis, and breast cancer mortality. Our data suggest that women who fail to undergo mammography are more often diagnosed with Stage II or greater breast cancer and are at increased risk of dying from breast cancer. Moreover, breast cancer was an important cause of death; breast cancer was the cause in 38% of all deaths in our study sample. These data support the use of regular mammography in older women and suggest that mammography can reduce breast cancer mortality for older women, even for women age 85 and older.

ACKNOWLEDGMENT

This study used the Linked Medicare-Tumor Registry Database. The interpretation and reporting of these data are the sole responsibility of the authors. The authors acknowledge the efforts of several groups responsible for the creation and dissemination of the Linked Database, including the Applied Research Branch, Division of Cancer Prevention and Population Science, National Cancer Institute; the Office of Information Services, and the Office of Strategic Planning, Health Care Financing Administration; Information Management Services (IMS), Inc.; and the Surveillance, Epidemiology, and End Results (SEER) Program Tumor Registries.

REFERENCES

1. Parker SL, Tong T, Bolden S, Wingo PW. Cancer statistics, 1997. CA Cancer J Clin 1997;47:5–27.
2. Yancik R, Ries LG, Yates JW. Breast cancer and aging women. A population-based study of contrasts in stage, surgery and survival. Cancer 1989;63:976–981.
3. Ries LAG, Kosary CL, Hankey BF et al., eds. SEER Cancer Statistics Review, 1973–1994. Bethesda, MD: National Cancer Institute, 1997. NIH Publication 97–2789.
4. Satariano WA, Belle SH, Swanson GM. The severity of breast cancer at diagnosis. A comparison of age and extent of disease in black and white women. Am J Public Health 1986;76:779–782.
5. Holmes FF, Hearne E. Cancer stage-to-age relationship. Implications for cancer screening in the elderly. J Am Geriatr Soc 1981;29:55–57.
6. Goodwin JS, Samet JM, Key CR et al. Stage at diagnosis of cancer varies with the age of the patient. J Am Geriatr Soc 1986;34:20–26.
7. Use of mammography services by women aged ≥ 65 years enrolled in Medicare—United States, 1991–1993. MMWR Morb Mortal Wkly Rep 1995;44: 777–781.
8. Shapiro S, Venet W, Strax P et al. Ten-to-fourteen year effect of screening on breast cancer mortality. J Natl Cancer Inst 1982;69:349–355.
9. Andersson I, Aspgren K, Janzon L et al. Mammographic screening and mortality from breast cancer: The Malmo mammographic screening trial. BMJ 1988;297:943–948.
10. Baker LH. Breast Cancer Detection Demonstration Project. Five-year summary report. CA Cancer J Clin 1982;32:194–225.
11. Roberts MM, Alexander FE, Anderson TJ et al. Edinburgh trial of screening for breast cancer. Mortality of seven years. Lancet 1990:335:241–246.
12. Tabar L, Gad A, Holmberg LH et al Reduction in mortality from breast cancer after mass screening with mammography. Randomised trial from the Breast Cancer Working Group of the Swedish National Board of Health and Welfare. Lancet 1985;1(8433):829–832.
13. Kopans DB Screening mammography in women over age 65. J Gerontol 1992;47:59–62.
14. Potosky AL, Riley GF, Lubitz JD et al. Potential for cancer related health services research using a Linked Medicare-Tumor Registry Database. Med Care 1993;31:732–748.
15. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative database. J Clin Epidemiol 1992;45:613–619.
16. SAS Institute. SAS/STAT User's Guide, Version 6, 4th Ed., Vol. 1 and 2. Cary, NC: SAS Institute, 1992.
17. Hosmer DW, Lemeshow S. Applied Logistic Regression. New York: John Wiley & Sons, 1989.
18. McCarthy EP, Burns RB, Coughlin SS et al. Mammography use helps to explain differences in breast cancer stage at diagnosis between older black and white women. Ann Intern Med 1998;128:729–736.
19. Lee ET. Statistical Methods for Survival Data Analysis, 2nd Ed. New York: John Wiley & Sons, 1992.
20. Tabar L, Fagerberg G, Chen HH et al. Efficacy of breast cancer screening by age. New results from the Swedish Two-County Trial. Cancer 1995;75:2507–2517.
21. Brown JT, Hulka BS. Screening mammography in the elderly. A case-control study. J Gen Intern Med 1988;3:126–131.
22. Faulk RM, Sickles EA, Sollitto RA et al. Clinical efficacy of mammographic screening in the elderly. Radiology 1995;194:193–197.
23. Day NE. Quantitative approach to the evaluation of screening programs World J Surg 1989;12:3–8.
24. Tabar L, Gad A, Holmberg L, Ljungquist U. Significant reduction in advanced breast cancer. Results of the first seven years of mammography screening in Kopparberg, Sweden. Diagn Imag Clin Med 1985;54:158–164.
25. Tabar L, Fagerberg G, Day NE, Holmberg L. What is the optimum interval between mammographic screening examinations? An analysis based on the latest results of the Swedish two-county breast cancer screening trial. Br J Cancer 1987;55:547–551.
26. Morrison AS, Brisson J, Khalid N. Breast cancer incidence and mortality in the Breast Cancer Detection Demonstration Project. J Natl Cancer Inst 1988;80:1540–1547.
27. Verbeek ALM, Holland R, Sturmans F et al. Reduction in breast cancer mortality through mass screening with modern mammography. First results of the Nijmegen Project, 1975–1981. Lancet 1984;8388:1222–1224.
28. Verbeek ALM, Holland R, Sturmans F et al. Mammographic screening and breast cancer. Age-specific effects in Nijmegen Project, 1975–1982. Lancet 1985;8433:865–866.
29. Van Dijck JA, Verbeek AL, Beex LV et al. Mammographic screening after age 65 years. Evidence for a reduction in breast cancer mortality. Int J Cancer 1996;66:727–731.

30. Blustein J. Medicare coverage, supplemental insurance, and the use of mammography by older women. N Engl J Med 1995;332:1138–1143.
31. Burns RB, McCarthy EP, Freund KM et al. Variability in mammography use among older women. J Am Geriatr Soc 1996;44:922–926.
32. Burns RB, McCarthy EP, Freund KM et al. Black women receive less mammography even with similar use of primary care. Ann Intern Med 1996;125: 173–182.
33. Bureau of Data Management and Strategy, Health Care Financing Administration. 1997 Data Compendium. Baltimore, MD: US Department of Health and Human Services, Health Care Financing Administration, 1997. Publication no. 03395.
34. Walter SD, Stitt LW. Evaluating the survival of cancer cases detected by screening. Stat Med 1987;6:885–900.
35. Costanza ME. Breast cancer screening in older women. Synopsis of a forum. Cancer 1992;69:1925–1931.

Comparing Standard Regression, Propensity Score Matching, and Instrumental Variables Methods for Determining the Effectiveness of Mammography in Older Women

Michael A. Posner, M.S.

Arlene S. Ash, Ph.D.

Michael Shwartz, Ph.D.

Karen M. Freund, M.D., M.P.H.

Mark A. Moskowitz, M.D.

Keywords –Selection Bias, Observational Studies, Health Services Research

Corresponding Author: Michael A. Posner, c/o Boston University School of Public Health, 715 Albany Street T4E, Boston, MA 02118. mposner@bu.edu. (617) 638-5873.

Grant support for this research was provided by the United States Army Medical Research and Materiel Command Breast Cancer Program, Grant #DMAD-17-97-1-7121.

Acknowledgements – This study used the linked Medicare-Tumor Registry database. The interpretation and reporting of these data are the sole responsibility of the authors. The authors acknowledge the efforts of several groups responsible for the creation and dissemination of the linked database, including the Applied Research Branch, Division of Cancer Prevention and Population Science, National Cancer Institute; the Office of Information Services and the Office of Strategic Planning, Health Care Financing Administration; Information Management Services, Inc.; and the Surveillance, Epidemiology, and End Results (SEER) Program tumor registries. In addition, the authors would like to thank Mingshan Lu and Ellen McCarthy for coding and Michael Ash for his economics expertise.

Abstract

In situations where randomized trials are not feasible, analysis of observational data must suffice. However, when using observational data, there is often selection bias for which we must account in order to adjust for pre-treatment differences between groups in their baseline characteristics. As an example of this, we used the Linked Medicare-Tumor Registry Database created by the National Cancer Institute and the Health Care Financing Administration to look at screening with mammography in older women to determine its effectiveness in detecting cancer at an earlier stage. The standard regression method and two methods of adjusting for selection bias are compared. We start with the standard analysis, a logistic regression predicting stage at diagnosis that includes as independent variables a set of covariates to adjust for differences in baseline risk plus an indicator variable for whether the person used screening. Next, we employ propensity score matching, which leads to an analysis which is more robust to model mis-specification than the standard analysis. Lastly, an instrumental variable analysis is conducted, which estimates the impact in the presence of unmeasured differences between the regular user and non-user group. This article reviews the assumptions involved in each of the analyses and compares the results.

Background

Randomized trials are viewed as the "gold standard" in research. However, there are many cases where these types of trials cannot be conducted, whether due to financial restrictions or ethical constraints. Where randomized trials are not feasible, observational data provide an excellent source of empirical information to address important clinical questions. Administrative databases are an important source of data for such observational studies, given their potential inclusion of large populations and the breadth of clinical and non-clinical data available. However, treatment and control groups potentially differ in their baseline characteristics, which makes the results susceptible to

bias. Only by adjusting for these differences in baseline characteristics can we determine the true treatment effect.

The standard method for analysis of a dichotomous outcome is a logistic regression. The exposure of interest is included as a predictor of outcome, with other covariates included to control for baseline differences. The odds ratio is then estimated, and tested to determine if it is statistically different from 1, which indicates no effect. This method will work if either all relevant baseline risk factors have been measured or the unmeasured factors are highly correlated with those that are measured, and if the relationship between risk factors and outcomes are correctly specified in the model. However, it is almost never reasonable to believe that all the important risk factors have been measured and it is often difficult to know whether relationships between measured risk factors and the outcomes have been correctly specified.

One method of adjustment that handles possible model mis-specification is the use of propensity score matching [9,18]. This process involves calculating for each case a propensity score, which indicates the likelihood (or "propensity") that each case is in the exposed group. Then, a sample of the exposed and unexposed groups is taken matching on the propensity score. As shown by Rosenbaum and Rubin, matching on the propensity score will result in a sub-sample of the study and control group that are well balanced in terms of observed risk factors. The standard analysis approach is then applied to the sub-sample data to determine the effect of the exposure on the outcome, after adjusting for any residual differences in observed covariates that remain after matching.

The third method that we employ is an instrumental variable approach. This approach is commonplace in econometrics [7], dating back to the 1920s [23], and is becoming increasingly popular in the health services outcomes research field as well [13,17,20]. For a variable, or group of variables, to be considered a good instrument, it should be neither associated with the outcome beyond its effect on exposure, nor with unmeasured confounders, after adjusting for those factors already in the model [8].

The use of mammography for screening women over age 70 years is an example of one of those areas where data are required to answer a critical clinical question, yet no

randomized trial information is available or likely to be available. It would be unethical to select a group of women to abstain from receiving screening mammography. While the data from most randomized controlled trials on screening mammography on women ages 50 to 70 years demonstrate a benefit of this procedure [1,16,19,22], there are no data to guide clinicians for women over age 70 years. Most of the trials included no women over age 70 years, and none of the trials reported any age-specific data within the 50 - 70 year old age groups to assess for any age related trends. Thus, the value of continuing screening in this group is yet to be established. Breast cancer incidence continues to rise beyond age 65, and accounts for 48% of all new breast cancers [15]. In the absence of data, clinicians are nevertheless making decisions regarding screening, and in some geographic areas, as many as one quarter of older women receive regular mammography [5]. A methodology to understand the benefits of screening mammography in women over age 70 years is critical.

Data Description

The database we utilized for this cohort study is the Linked Medicare-Tumor Registry Database. The linked database was jointly created by the National Cancer Institute (NCI) and the Health Care Financing Administration (HCFA) [14]. The database links Medicare data on women ages 65 and older from 1985 to 1994 with cancer registry information from the NCI's SEER program for cancers diagnosed between 1973 and 1993. The two databases overlap in three racially and socially diverse geographic areas: metropolitan Atlanta, Seattle-Puget Sound, and the state of Connecticut.

Medicare Physicians' Claims files, which provide a record for every physician claim covered under Medicare Part B, were employed to measure mammography utilization and primary care utilization. Medicare MedPAR (Medical Provider Analysis and Review) files, which provide a record for every inpatient hospitalization, were utilized to develop measures of comorbidity. Medicare beneficiary enrollment files were used to determine race. The SEER information was used to determine age at diagnosis,

stage at diagnosis using the Tumor, Nodal Status, Metastases (TNM) classification system [3], geographic location, and cause of death. 1990 U.S. Census data were used to obtain median household income by zip code, which was used as an ecological measure of socioeconomic status.

Our study sample consisted of all women with a first diagnosis of breast cancer in the three geographic areas for whom we could track 2 years prior to the diagnosis of breast cancer [10]. Since the Medicare utilization files provided mammography data on women aged 65 and older beginning in 1985, our sample included women age 67 and older with their diagnosis in 1987 or later. We excluded women without two full years of Medicare Part B claims data prior to their diagnosis; this meant we excluded women with health maintenance organization coverage since no part B claims are filed for these women.

We utilized the following procedure to classify mammography use of women. Women were classified as regular mammography users if they had claims for two separate bilateral mammograms (CPT code 76091 or 76092) within the two years prior to their breast cancer diagnosis, which were at least 10 months apart. Non-users were those women with no mammography claims during the two years prior to their diagnosis. Those women with less frequent mammography or mammography prior to diagnosis comprise a heterogeneous group of women receiving both screening and diagnostic studies and were excluded from further analyses. Stage at diagnosis, our primary outcome variable, was classified as early (in situ and Stage I) or late (Stage II, III, and IV). Women with unstaged cancer were excluded from further analysis. Our previous work indicates that this group of women has both early and late stage disease, as demonstrated by a survival rate intermediate between the early and late stage groups [10].

The Linked database contains a number of clinical and non-clinical variables that are likely to be predictive of the outcome, and therefore were utilized as covariates in each model. Age at diagnosis was considered as both a continuous and categorical variable (67-69, 70-74, 75-69, 80-84 and 85+ years). Final analysis used the categorical listing of age. Comorbidity was measured using a modified Charlson Cormobidity Index from inpatient claims' diagnosis codes [6]. Women were categorized as having no

hospital inpatient claims (i.e. comorbidity could not be assessed), or being hospitalized with either none or one or more comorbid conditions. Race was recorded as black or non-black (white, hispanic, other). Median household income of zip code was considered both as a continuous variable and by quintile for each of the three geographic regions. After looking at quintiles by region, the variable used in analysis was a dichotomized split of the highest 40% versus the lower 60%. Number of claims for office visits to primary care providers prior to breast cancer diagnosis were considered both continuously and categorically (0, 1-3, 4-12, 13+). The final analysis used the categorical approach.

Methods

For the standard analysis, a logistic regression was run to predict stage at diagnosis from user status, controlling for age, race, comorbidity, regional income, and primary care visits, which were included as covariates in the model. The c-statistic was determined as a measure of predictive validity.

For the propensity score approach, a logistic regression model including age, race, comorbidity, regional income, region, and primary care visits was used to determine the propensity of being a user. A split into deciles based on the propensity scores was done. To develop the matched samples, within each decile a random sample of the larger group (regular users or non-users) was taken in order to get the same number as was in the smaller group (see table 1). The reduced samples were then combined into the analytic dataset. A second model excluding region was also calculated to determine if the specification of propensity differed by region. The results did not differ so this alternate method of generating propensity scores was not pursued further. To examine the extent to which the above matching resulted in samples of regular users and non-users more comparable in terms of baseline characteristics, p-values for chi-squared tests of independence were calculated for the categorical risk factors variables and p-values for independent sample t-tests for equivalence of population means were calculated for numeric risk factors.

For the instrumental variable analysis, the first step is to determine which variable or variables are to be used as instruments. In our situation, a candidate for an instrument must be a predictor of whether someone is a regular user of mammography with no residual predictive power on stage at diagnosis, after controlling for the other covariates in the model, including the propensity of being a regular user. Angrist, Imbens, and Rubin (1996) describe how this can be broken down into two necessary conditions. We consider these conditions in the context of our example, using region as our instrument. The first necessary condition is that the effect of moving from one region to another, conditional on user status and the measured covariates, does not change the expected outcome. For example, we would expect that a woman with certain characteristics (primary care visits, age, race, etc.) receiving regular screening in Seattle, would have the same likelihood of early stage disease diagnosed from mammography had she moved to Atlanta or Connecticut. If this assumption were not met, it would imply that, after conditioning on observed covariates, follow-up after a positive mammogram in one region is different than follow-up in another region. There is no evidence to suggest this is the case. The second condition is that there is random, or at least ignorable, assignment to whether or not the woman was a regular user. Failure to meet this assumption would imply that there are unmeasured variables that differ across regions that are associated with race, age, primary care visits (which might be thought of as a proxy for access), income, or comorbidities. These covariates are also associated with whether or not someone becomes a regular user. Though there is no way to empirically validate this last assumption, it seems reasonable in the context of our example.

The instrumental variable approach involves a two-stage model. In the first stage, covariates plus the instruments are used to predict user status. The predicted probability of being a regular user is then used in lieu of user status as an independent variable in the second stage, along with any measured covariates. Variables used as instruments in the first stage model are excluded, since these variables are assumed to effect the outcome only through their association with user status. The coefficient associated with predicted user is a measure of the impact of use.

As a measure the strength of the instrumental variables, we used results from the first stage model. The odds ratio as well as the wald chi-squared value was used. Staiger and Stock (1997) suggest that a partial F-statistic of about 10 is sufficient to not be a weak instrument. Instead we use Wald chi-squared statistics from our logistic regression models, calculated as the difference in the –2 log likelihoods between the full and reduced (excluding the instrument) models to predict user status. The difference between the partial F (square of a t-distribution) and Wald chi-squared (square of a normal distribution) should be insignificant due to our sample size, so that the same rule of thumb may apply. In addition, the odds ratio of each group is presented as another measure of the strength of the association.

Results

The final sample size was 4667 women. Of the 4667 women, 11 were excluded from analyses due to missing zip codes, which made the zip code-income matching impossible. Of the remaining 4656, 1354 were diagnosed with late stage cancer and 3302 with early stage cancer; 2516 were regular users and 2140 were non-users.

The standard model controlling for age at diagnosis, race, comorbidities, regional income, region, and number of primary care had a c-statistic of 0.68. The conclusion, based on this model, is that regular users have 2.97 times the odds of being diagnosed at early stage relative to non-users (95% CI: 2.56, 3.45). These results are similar to those found by McCarthy (2000).

In the propensity score analysis, the model to predict user status using age, race, comorbidities, regional income, region, and primary care visits as independent variables had a c-statistic of 0.78. Table 1 shows the pre- and post-sampling number of cases by decile and compares baseline characteristics of regular users and non-users.

We can see that the matching resulted in much more balance between the groups in terms of the measured covariates. Each of these variables showed a statistical difference based on user status before sampling yet no difference after sampling. The exception to this is stage at diagnosis, which was not included in the propensity score

model, since it occurs after use status and hence could not influence the likelihood of being a regular user.

When the logistic regression was run on the matched sample, the odds ratio associated with user status was 3.27 (95% CI: 2.72, 3.93), a result similar to the pre-sampling results from the standard logistic regression analysis.

The result from the instrumental variable approach using region as an instrument produced an odds ratio of 3.01 (95% CI: 1.09, 8.34), again similar to the standard analysis. The chi-squared statistic for not using region as a predictor of user was 53.0 ($p<.001$). The odds ratio for predicting user status from region is 1.34 for Seattle and 0.64 for Atlanta, both using Connecticut as the reference group.

In addition, we considered two other instruments. First was race (black vs. other). The odds ratio for early detection for regular users relative to non-users is 3.69 (95% CI: 0.88, 15.57). The chi-squared statistic for not using race as a predictor of user was 5.74 ($p=.017$). The odds ratio for predicting user status from race is 0.67, indicating that blacks have two-thirds the odds of non-blacks to become a regular user of mammography.

The second alternative instrumental variable analysis we conducted was using primary care visits as the instrument. The resulting odds ratio is 2.45 (95% CI: 1.80, 3.35). The chi-squared statistic for not using primary care visits as a predictor of user was 824 ($p<.001$). The odds ratio for predicting user status from number of primary care visits is 3.87, 9.73, and 14.70 for 1-3, 4-12, and more than 12 primary care visits, respectively, all relative to those women with no primary care visits.

Discussion

The standard analysis gives us results to which we can compare the others in order to determine the amount of model mis-specification (when results are compared to the propensity score results) and the effect of unmeasured confounders (when the results are compared to the instrumental variable analysis). It is important to note that the odds ratio

of 2.97 from the standard analysis is a measure of effect of use for women of a type similar to those in the entire sample.

The propensity score matching method produced an odds ratio of 3.27. While this effect is slightly larger than that from the standard analysis, the difference is clearly not statistically significant, based on the confidence intervals. The propensity score matching approach is an estimate of the impact of being a regular user of mammography for the sample that is similar in terms of the measured covariates that were included in the propensity generation model. Since this result is so close to that of the standard model, it implies that there is not a large bias that has come into effect in the standard model due to a biased selection of women who have self-selected to be regular users or non-users of mammography.

The main instrument under consideration was region. This variable meets the conditions of unchanged outcome resulting from changing regions as well as ignorable assignment to treatment, as stated in the methods section. In addition, the significance of region on predicting user status was quite strong (Chi-squared of 53, $p<.001$ from the reduced model), indicating that we do not have a weak instrument. For otherwise identical women with average characteristics, the effect of being in Seattle rather than Connecticut increases the odds of mammogram use by 34%. This is consistent with shown geographic variation in practice patterns, in particular those of breast cancer care [11,12]. A weak instrument results in a bias in estimating program impact toward the estimate from the standard approach. If our instrument were weak, we could not conclude that similarity of instrumental variable estimates and standard analysis estimates implied that there were not unmeasured confounders.

The instrumental variable analysis, using region as the instrument, produced an odds ratio of 3.01. Again, this result is similar to that of the standard analysis, implying that there is no large bias coming from unmeasured variables. This value measures the impact among those whose behavior changes depending on their region, with all else being equal. The confidence interval for this value (1.09, 8.34) is much wider than that of the standard model (2.56, 3.45). This is expected, because the instrumental variable approach depends on much less variation in user status, i.e., exclusively on differences in

user status caused by regional differences, than do the alternative models. With less variation in user status, the estimate of the standard error will be larger.

We considered two alternate instruments to compare their results to those found in the primary model. The first was race. Race was chosen, since it was correlated with region (correlation coefficients of .25, -.12, and -.08 with Atlanta, Seattle, and Connecticut, respectively). However, it was assumed that race does have a residual effect on outcome beyond its effect on user. This violates the first assumption of being an instrument. In addition, race produced a chi-squared value of 5.74, substantially less than the value of 10 recommended by Staiger and Stock (1997). Thus, even if race were accepted as an instrument from a theoretical argument, it would be a weak one, thus invalidating its use as an instrument.

The second alternate instrument considered was number of primary care visits (as a categorical variable). Primary care visits, however, violates the first assumption of robust effects from changing values of the instrument. It seems reasonable, for example, that changing someone from no primary care visits to 13 or more primary care visits will likely increase the probability of their cancer being detected early beyond its effect on whether or not the woman was a regular user. While the chi-squared value for primary care visits predicting user was strong (824 with 3df), the result is irrelevant, since this variable violates the initial assumptions necessary to be considered as an instrument.

In summary, all three analyses, the standard regression, the propensity score matching, and the instrumental variable analysis using region as the instrument, produced very similar results. The conclusion from this is that the standard regression analysis is valid. There is little model mis-specification, either from measured variables, as seen via the propensity score matching, or from unmeasured variables, as seen via the instrumental variable analysis. The use of an instrument must satisfy both of the conditions set out by Angrist, Imbens, and Rubin (1996) of unchanged outcome resulting from changing regions as well as ignorable (or random) assignment to treatment as well as be a strong predictor. The rule of thumb for being a strong predictor, according to Staiger and Stock (1997), is a partial F statistic of at least 10 in its prediction of exposure. We propose here that the use of a reduced model Chi-squared in the context of a logistic regression is

equivalent with large sample analysis. Once a variable is shown to be consistent with these assumptions, it may be considered as an instrument in the analysis and the impact of the exposure on the outcome can be compared to other models to determine the effect on selection bias on the results.

Table 1: Propensity Score Matching Results						
	Pre-Sampling			Post-Sampling		
	Non-User	User	p-value	Non-User	User	p-value
Total Sample	2140	2516		1274	1274	
Decile 1	416	**57**		57	**57**	
Decile 2	339	**89**		89	**89**	
Decile 3	359	**136**		136	**136**	
Decile 4	239	**205**		205	**205**	
Decile 5	**204**	305		**204**	204	
Decile 6	**158**	287		**158**	158	
Decile 7	**135**	321		**135**	135	
Decile 8	**112**	366		**112**	112	
Decile 9	**100**	379		**100**	100	
Decile 10	**78**	371		**78**	78	
Age at Diagnosis	77.2	74.5	0.001	75.8	75.6	0.428
Age at Diagnosis						
67-69	37.8%	62.2%		48.9%	51.1%	
70-74	39.1%	60.9%		51.4%	48.6%	
75-79	44.1%	55.9%		49.7%	50.3%	
80-84	51.6%	48.4%		49.1%	50.9%	
85+	76.5%	23.5%	0.001	50.0%	50.0%	0.919
Charlson Comorbidities						
Not Hospitalized	43.6%	56.4%		48.7%	51.3%	
Hosp, No Comorbidities	43.0%	57.0%		50.3%	49.7%	
At Least One Comorbidity	56.9%	43.1%	0.001	51.0%	49.0%	0.699
Race						
Black	63.2%	36.8%		55.9%	44.1%	
Non-Black	45.1%	54.9%	0.001	49.7%	50.3%	0.143
Income (Median of Zip Code)	$42,030	$41,137	0.061	$42,004	$41,727	0.672
Regional Income						
Top 40%	44.1%	55.9%		50.4%	49.6%	
Lower 60%	47.3%	52.7%	0.036	49.7%	50.3%	0.748
Primary Care Visits	4.9	10.5	0.001	7.6	8.2	0.074
Primary Care Visits						
None	83.1%	16.9%		51.2%	48.8%	
1-3	55.7%	44.3%		50.0%	50.0%	
4-12	32.8%	67.2%		50.1%	49.9%	
13+	26.5%	73.5%	0.001	49.2%	50.8%	0.952
Location						
Seattle	35.8%	64.2%		49.6%	50.4%	
Atlanta	55.0%	45.0%		49.5%	50.5%	
Connecticut	50.0%	50.0%	0.001	50.5%	49.5%	0.908
Stage						
Early (TNM 0 or I)	38.2%	61.8%		41.6%	58.4%	
Late (TNM II, III, or IV)	64.8%	35.2%	0.001	69.6%	30.4%	0.001

Affiliation/Email addresses of Authors:

Michael A. Posner[1]	mposner@bu.edu
Arlene Ash[2]	aash@bu.edu
Michael Shwartz[3]	mshwartz@bu.edu
Karen Freund[2]	karen.freund@bmc.org
Mark Moskowitz[2]	mark.moskowitz@bmc.org

1 *Data Coordinating Center, Boston University School of Public Health, Boston, Massachusetts.*

2 *Health Care Research Unit, Section of General Internal Medicine, Evans Department of Medicine, Boston Medical Center, Boston, Massachusetts.*

3 *Operations Management, Boston University School of Management, Boston, Massachusetts.*

References

1. Andersson I, Aspgren K, Janzon L et al. Mammographic Screening and Mortality from Breast Cancer: The Malmo Mammographic Screening Trial. BMJ 1998;297:943-948.

2. Angrist JD, Imbens, GW, Rubin, DB. Identification of Causal Effects Using Instrumental Variables. JASA 1996;96:444-72.

3. Beahers OH, Myers MH. Manual for Staging of Cancer. 2nd edition. Philadelphia: JB Lipincott, 1983.

4. Bound J, Jaeger DA, Baker RM. Problems with Instrumental Variables Estimation When the Correction Between the Instruments and the Endogenous Explanatory Variable is Weak. JASA 1995;90:443-50.

5. Burns RB, McCarthy, Freund KM, et al. Variability in Mammography Use Among Older Women. JAGS 1996;44:922-926.

6. Deyo RA, Cherkin DC, Ciol MA Adapting a Clinical Comoridity Index for Use with IDC-9-CM Administrative Database. J Clin Epidemiol 1992;45:613-19.

7. Greene WH. Econometric Analysis. Prentice Hall, New Jersey. 1997. 288-95.

8. Greenland S. An Introduction to Istrumental Variables for Epidemiologists. Int'l J Epi 2000;29:722-9.

9. Joffe MM, Rosenbaum PR. Invited Commentary: Propensity Scores. Am J Epidemiol 1999;150:327-33.

10. McCarthy EP, Burns RB, Freund KM, Ash AS, Shwartz M, Marwill SL, Moskowitz MA. Mammography Use, Breast Cancer Stage at Diagnosis, and Survival Among Older Women. J Am Geriatr Soc 2000;48:1226-33.

11. Nattinger AB, Gottlieb MS, Veum J, Yagnke DL, Goodwin JS. Goegraphic variation in the use of brest conserving treatment for breast cancer. N Engl J Med. 1992;326:1102-1107.

12. Nattinger AB, Gottlieb MS, Hoffman RG, Walker AP, Goodwin JS. Minimal increase in use of breast-conserving surgery from 1986 to 1990. Med Care. 1996;24:479-489.

13. Newhouse JP, McClellan M. Econometrics in Outcomes Research: The Use of Instrumental Variables. Ann Rev Public Health 1998:19:17-34.

14. Potosky AL, Riley GF, Lubitz JD, Mentnech RM, Kessler LG. Potential for cancer realted health services research using a Linked Medicare-Tumor Registry Database. Med Care 1993;31:732-748.

15. Ries, LAG, Kosary CL, Hankey BF, et al., eds. SEER Cancer Statistics Review, 1973-1994. Bethesda, MD: National Cancer Institute, 1997. NIH Publication 97-2789.

16. Roberts M, Alexander FE, Anderson TJ et al. Edinburgh Trial of Screening for Breast Cancer Mortality of Seven Years. Lancet 1990;335:241-246.

17. Robins JM. The Analysis of Randomized and Non-Randomized AIDS Treatment Trials Using a New Approach to Causal Inference in Longitudinal Studies. In: Sechrest L, Freeman H, Mulley A (eds). Health Service Research Methodology: A Focus on AIDS. Washington, DC: US Public Health Service, 1989, pp.113-59.

18. Rosenbaum PR, Rubin DB. The Central Role of the Propensity Score in Observational Studies for Causal Effects. Biometrika 1983;70:41-55.

19. Shapiro S, Venet W, Strax P et al. Ten to Fourteen Year Effect of Screening on Breast Cancer Mortality. J Natl Cancer Inst 1982;69:349-355.

20. Sheiner LB, Rubin DB. Intention-to-Treat Analysis and the Goals of Clinical Trials. Clin Pharm Therap 1995;57:6-15.

21. Staiger D, Stock JH. Instrumental Variables Regression with Weak Instruments. Econometricia 1997;63;3;557-86.

22. Tabar L, Gad A, Holmberg LH, et al. Reduction in Mortality from Breast Bancer After Mass Screening with Mammography. Randomized Trial from the Breast Cancer Working Group of Swedish National Board of Health and Welfare. Lancet 1985;1:829-832.

23. Wright S. Appendix. In: Wright PG (ed.). The Tariff on Animal and Vegetable Oils. New York: Macmillan, 1928.